



Máster Online en Buscadores

Selección de
unidades didácticas
2007/2008

*Cristòfol Rovira
Lluís Codina
Mari-Carmen Marcos
Rafael Pedraza*

*Àrea Docente de Ciències de la Documentació
Departament de Comunicació
Universitat Pompeu Fabra*

www.masterenbuscadores.com

Dirección editorial: Cristòfol Rovira
Diseño del anagrama de la portada: Ignacio Pastor

Máster Online en Buscadores. Selección de unidades didácticas 2007/2008
Cristòfol Rovira , Lluís Codina , Mari-Carmen Marcos y Rafael Pedraza
1ª edición, septiembre 2008

Editado por los autores bajo la licencia Creative Commons del tipo:
Reconocimiento-No comercial-Sin obras derivadas 2.5 España
<http://creativecommons.org/licenses/by-nc-nd/2.5/es/>

© Cristòfol Rovira , Lluís Codina , Mari-Carmen Marcos y Rafael Pedraza, 2008
ISBN: en trámite
DL: en trámite



Introducción

Este documento es una selección del texto completo 6 unidades didácticas del Master Online en Buscadores impartido durante el curso 2007/2008.

El Máster en Buscadores es una iniciativa del grupo de investigación DigiDoc del Área Docente de Ciencias de la Documentación del Departamento de Comunicación de la Universidad Pompeu Fabra. Está organizado por el Instituto de Educación Continua de la Universidad Pompeu Fabra bajo la dirección de Cristòfol Rovira, la codirección de Lluís Codina y la coordinación académica de Mari-Carmen Marcos.

Los contenidos del Máster en Buscadores se imparten de forma no presencial y a través de Internet entre octubre y junio. Los contenidos del máster están formados por 72 unidades didácticas organizadas en los siguientes 6 módulos:

- Módulo F: Fundamentos en buscadores
- Módulo P: Posicionamiento en Buscadores (SEO)
- Módulo M: Márketing en buscadores (SEM)
- Módulo V: Vigilancia tecnológica
- Módulo R: Recuperación de la información
- Módulo E: Buscadores especializados

Este documento contiene el texto completo de una unidad didáctica representativa de cada módulo, a menudo la unidad inicial, en la cual se ofrece una panorámica general de todos los contenidos del módulo que posteriormente serán desarrollados por otras unidades. Los textos fueron actualizados por última vez en junio del 2008.

Este documento nace con la intención de dar a conocer los contenidos del máster para su difusión entre los profesionales del sector.

Sobre el grupo de investigación Digidoc

- Grupo DigiDoc <http://www.digidocweb.net>
- Publicación anuario Hipertext.net <http://www.hipertext.net>
- Línea de investigación Web semántica <http://www.semanticweb.net>
- WebLog profesional Servicio de Alerta <http://docdigital.typepad.com>
- Máster en Buscadores <http://www.masterenbuscadores.com>
- Máster en Documentación Digital <http://www.documentaciondigital.org>

Sobre los autores

Dr. Cristòfol Rovira
Universidad Pompeu Fabra
<http://www.cristofolrovira.com>



Dr, Lluís Codina
Universidad Pompeu Fabra
<http://www.lluiscodina.com>



Dra. Mari-Carmen Marcos
Universidad Pompeu Fabra
<http://www.mcmarcos.com>



Rafael Pedraza
Universidad Pompeu Fabra
<http://www.technobserver.com>



Sumario de unidades didácticas seleccionadas

Módulo F: Fundamentos en buscadores

Unidad F092 Estructura y funciones de un buscador por Lluís Codina

Módulo P: Posicionamiento en Buscadores (SEO)

Unidad P235 El ciclo de vida de una campaña de posicionamiento por Lluís Codina

Módulo M: Márketing en buscadores (SEM)

Unidad M519 Fundamentos del marketing en buscadores por Cristòfol Rovira

Módulo V: Vigilancia tecnológica

Unidad V446 Vigilancia tecnológica e inteligencia competitiva para SEM-SEO por Cristòfol Rovira

Módulo R: Recuperación de la información

Unidad R707 Buscadores y presentación de resultados en agrupación por Mari-Carmen Marcos

Módulo E: Buscadores especializados

Unidad E665 Buscadores de mapas por Rafael Pedraza

Sumario general del Máster en Buscadores

Módulo F: Fundamentos en buscadores

F 092 Estructura y funciones de un buscador

- F 910 Características básicas de buscadores y directorios
- F 911 Spiders de buscadores: características y funcionamiento
- F 094 Web semántica y servicios de búsqueda
- F 701 Estudio comparativo de los formularios de búsquedas avanzadas
- F 655 Análisis comparativo de los principales buscadores generalistas
- F 095 Algoritmos de posicionamiento en los buscadores
- F 091 Historia de los buscadores
- F 096 Recursos online sobre buscadores
- F 702 Ética y privacidad en buscadores
- F 914 entrevista Entrevista y debate con Fernando Maciá

Módulo P: Posicionamiento en Buscadores (SEO)

P 235 El ciclo de vida de una campaña de posicionamiento

- P 240 Factores críticos del posicionamiento en buscadores
- P 244 Optimización del código (X)HTML para mejorar el posicionamiento
- P 097 Internet invisible: allí donde no llegan los buscadores
- P 704 Alta en buscadores
- P 237 Obtención de enlaces entrantes
- P 239 Como evitar las penalizaciones de los buscadores: spam de buscadores
- P 241 Herramientas freeware para mejorar el posicionamiento en buscadores
- P 521 Como contratar un servicio de SEO a una empresa
- P 912 Buscadores y metadatos
- P 916 entrevista Entrevista y debate con Ismael El-Qudsi

Módulo M: Márketing en buscadores (SEM)

M 519 Fundamentos del marketing en buscadores

- M 703 Google Adwords: introducción y alta de una campaña
- M 531 Yahoo Sponsored Search: introducción y alta de una campaña
- M 532 Campañas en contenidos
- M 533 Informes y seguimiento de campañas
- M 522 Introducción a Google Analytics
- M 524 Inserción de publicidad en la propia web: Google Adsense
- M 527 Seguimiento de la presencia en buscadores: Google Webmaster Tools
- M 529 Rastreo automático: DigiDocSpider
- M 523 Definición de objetivos y cálculo del retorno de la inversión
- M 915 entrevista Entrevista y debate con Montserrat Peñarroya

Módulo V: Vigilancia tecnológica

V 446 Vigilancia tecnológica e inteligencia competitiva para SEM-SEO

- V 658 Metabuscadores y multibuscadores
- V 705 Buscadores especializados: visión panorámica
- V 668 Buscadores de noticias
- V 420 Evaluación de documentos digitales
- V 396 Sindicación de contenidos: RSS y ATOM
- V 623 (a) Minería de datos. (b) Búsqueda de patentes
- V 671 Internet Invisible
- V 382 Agentes de búsqueda
- V 449 (a) Monitorización. (b) Servicios de Alerta
- V 917 entrevista Entrevista y debate con Juan Carlos Vergara

Módulo R: Recuperación de la información

- R 93 Lenguajes de interrogación para buscadores
- R 398 Técnicas de visualización de información aplicables a buscadores
- R 707 Buscadores y presentación de resultados en agrupación (clustering)
- R 377 Búsquedas de búsquedas: Google Trends
- R 385 Web 2.0 y búsquedas: folksonomias
- R 390 Servicios online de indexación y búsqueda: FreeFind, PicoSearch
- R 397 Creación de buscadores especializados
- R 383 Personalización de los formularios de búsqueda
- R 387 Creación de buscadores online y offline con DigiDocIndex
- R 389 (a) Inconsistencias de los buscadores. (b) BusBas: buscador básico
- R 918 entrevista Entrevista y debate con Ricardo Baeza-Yates

Módulo E: Buscadores especializados

- E 376 Servicios de marcadores con etiquetas
- E 651 Buscadores de documentos audiovisuales
- E 652 Buscadores de Blogs
- E 653 Buscadores del texto de libros impresos
- E 654 Buscadores para usos académicos
- E 662 Buscadores con indexación manual e interrogación en lenguaje natural
- E 665 Buscadores de mapas
- E 666 Buscadores de informes y estudios de mercado
- E 657 Buscadores de sedes web antiguas: el archivo de Internet
- E 709 Buscadores de música
- E 710 Buscadores de software
- E 714 Buscadores de información personal
- E 656 Buscar en diccionarios, enciclopedias y obras de referencia
- E 708 Buscadores de precios
- E 664 Buscadores de cursos y empleo
- E 661 Buscadores por el ranking de visitas
- E 919 entrevista Entrevista y debate con José Ramón Pérez Agüera

Título	Artículo F092. Estructura y funciones de un motor de búsqueda
Autor	Lluís Codina
Palabras clave	Rastreadores, Sistema de indización, Índice directo, Índice invertido, Sistemas de recuperación de información, Página de resultados
Descripción	Las funciones principales de un motor de búsqueda se identifican con las de los sistema de información documental, a saber: satisfacer necesidades de información a partir del análisis de documentos. No obstante, la Web presenta características propias a las que deben dar respuesta los motores de búsqueda mediante unas estructuras y un conjunto de funciones bien determinados.
Objetivos	El principal objetivo de la unidad didáctica es presentar la estructura típica y las funciones más características de los motores de búsqueda, en tanto sistemas de representación y de recuperación de información.

Sumario

1	Introducción.....	2
2	Sistemas de información documental.....	2
3	Estructura de un motor de búsqueda.....	3
4	Rastreador del motor de búsqueda.....	4
5	Sistema de recuperación de información.....	4
5.1	Sistema de indización.....	4
5.1.1	Índices directos.....	5
5.1.2	Índices invertidos.....	6
6	Interfaz de consulta.....	7
6.1	Cálculo de relevancia.....	8
6.1.1	Criterios intrínsecos.....	9
6.1.2	Criterios externos.....	9
7	Página de resultados.....	9
8	Conclusiones.....	10
9	Bibliografía.....	10

1 Introducción

Los motores de búsqueda son los más importantes sistemas de información documental de nuestro tiempo. Forman parte del paisaje de Internet de una manera que parece definitiva. Nadie imagina hoy la posibilidad de explotar la inmensa riqueza de la Web sin el auxilio de los motores de búsqueda cuyas funciones no han dejado ni de aumentar ni de experimentar cambios desde su temprana aparición en los años 90.

2 Sistemas de información documental

Una forma de entender las funciones de un motor de búsqueda consiste en concebirlo como una clase de sistema de información documental, que se puede representar con el diagrama (Ilustración 1) que mostramos a continuación:

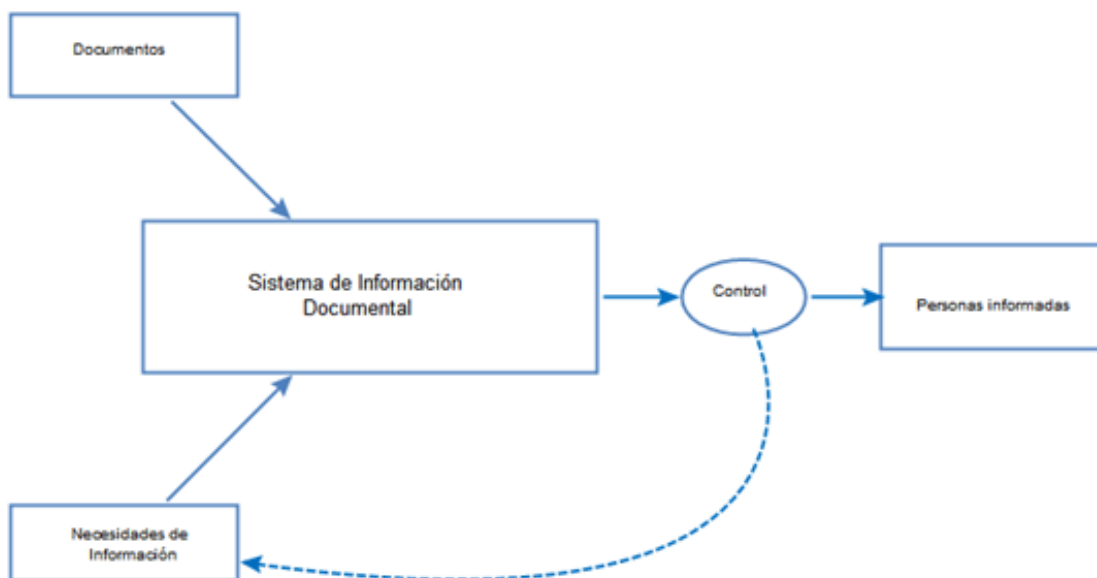


Ilustración 1: Diagrama de un Sistema de Información documental

Podemos ver que un sistema documental (tal como un motor de búsqueda) acepta dos clases de entradas: **Documentos** y **Necesidades de Información**. Los documentos en nuestro caso son páginas web principalmente y las necesidades de información son las preguntas de los internautas. La salida del sistema es una lista de sitios web que, presumiblemente, contienen información capaz de satisfacer las necesidades de información del usuario, de aquí que la salida se rotule con el (optimista) texto de **Personas informadas**. Naturalmente, si la primera consulta no da buenos resultados, lo lógico es que el usuario (el internauta) haga algún intento más. Esto queda representado por la elipse rotulada con la palabra **Control**. Es decir, se supone que el usuario

examina la relevancia de los resultados y si no le satisfacen, reitera la pregunta usando otros términos o introduciendo algún otro tipo de cambios.

3 Estructura de un motor de búsqueda

Un motor de búsqueda (MDB a partir de ahora) consiste en un conjunto de programas que trabajan en cooperación. Estos programas se pueden distribuir en dos grupos principales:

- *Sistema de exploración o Rastreador* (en inglés: *spyder* o *crawler*)
- *Sistema de recuperación de información* (en inglés: *information retrieval system*)

Cada uno de estos sistemas se compone de otros subsistemas (Ilustración 2) que analizaremos a continuación:

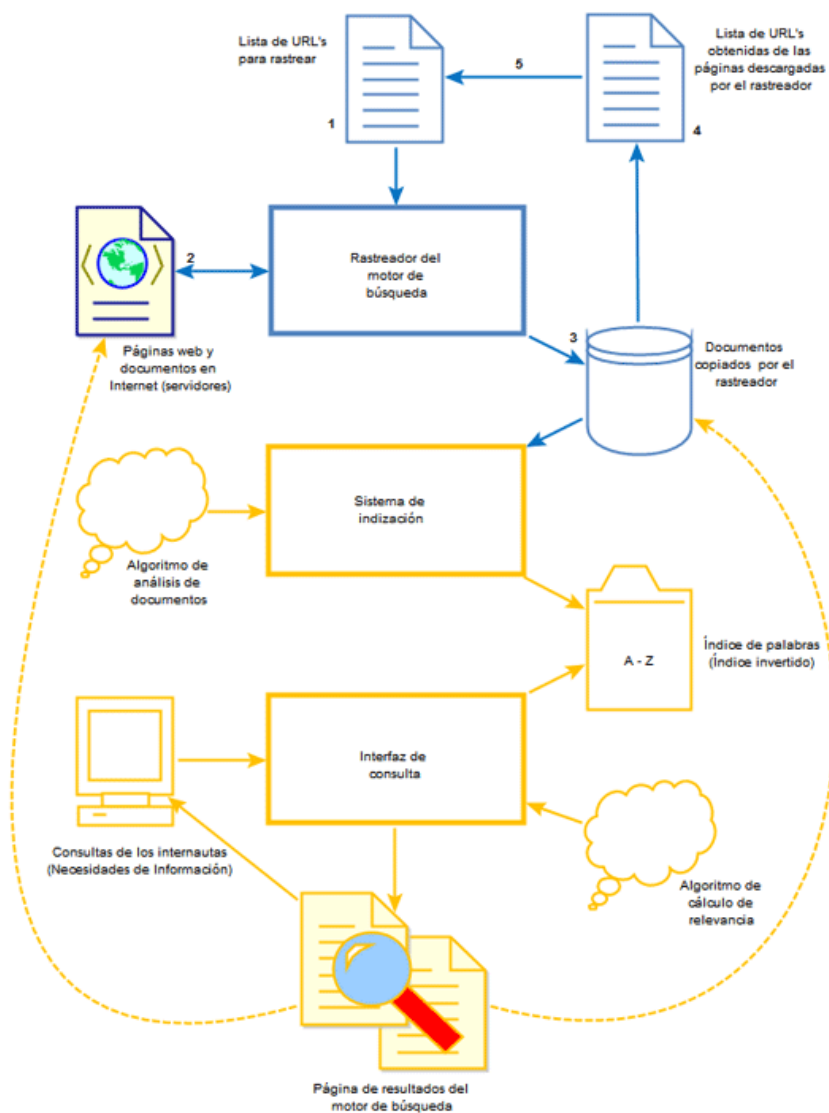


Ilustración 2: Estructura y funciones de un Motor de Búsqueda

4 Rastreador del motor de búsqueda

El rastreador, o *sistema de exploración* tiene la misión de descubrir y descargar páginas web y otros documentos de Internet. Para ello, parte de una lista de URL o direcciones de páginas web (1). A partir de ella accede a documentos (2) publicados en la web. Descarga esos documentos, en su mayoría páginas web, pero también documentos ofimáticos, presentaciones, archivos de imágenes, etc. en el almacén o repositorio de documentos del MDB (3). A partir de las páginas descargadas, el rastreador obtiene nuevas URL (4) que añade a su lista de URL pendientes de analizar (5).

Las tareas del rastreador son más complejas si se analizan con detalle, pero dado que tenemos una unidad dedicada a los rastreadores, en lo que sigue, nos concentraremos en el sistema de recuperación de información.

Actividad 1: Acuda a la versión inglesa de Wikipedia (http://en.wikipedia.org/wiki/Web_crawler) y haga una revisión del artículo, aunque sea en diagonal, prestando atención a los diagramas y a la lista de crawlers que presenta el artículo. *Alternativa* : Vaya a la versión española de la Wikipedia (http://es.wikipedia.org/wiki/Web_crawler) ¿Qué le parece el artículo examinado?. ¿Propondría algún cambio al mismo?

5 Sistema de recuperación de información

El sistema de recuperación de información tiene a su vez dos componentes:

- Sistema de indización
- Sistema de consulta

El primero de ellos se ocupa del análisis de los documentos descargados de la web y de la creación de los índices que permiten luego las consultas al motor; mientras que el segundo constituye la interfaz visible del MDB, precisamente la parte con la que interactúan los usuarios.

5.1 Sistema de indización

Si el MDB puede responder preguntas en el asombrosamente breve tiempo al que nos han acostumbrado los otros motores de búsqueda (típicamente, fracciones de segundo), es porque no exploran la web para los usuarios en tiempo real, es decir, en el momento de la consulta.

Un sistema que pretendiera explorar de forma secuencial los documentos de la web para ofrecer respuestas no tendría el más mínimo sentido. Podrían

transcurrir mese o años entre la pregunta y la respuesta. Si un estudiante hiciera su pregunta en primer curso de la carrera, cuando obtuviese la respuesta ya se habría licenciado (eso si no suspendió curso esperando la respuesta del motor). En lugar de lo anterior, lo que hacen los motores es consultar sus índices internos.

Actividad 2: Haga diversas consultas en Google o en Yahoo y compare los tiempos de respuestas (indicados en una barra situada justo sobre la lista de resultados). Compruebe si, con diferentes preguntas, por ejemplo con términos relativamente específicos (como "stradivarius") o con términos mucho más genéricos (como "música") los tiempos de respuesta son muy distintos o, en general, los tiempos son similares. Cabe esperar que no haya diferencias apreciables. El motivo se discutirá más adelante.

5.1.1 Índices directos

Los MDB utilizan al menos dos clases de índices, denominados respectivamente índices directos e índices invertidos. Las consultas las resuelven los *índices invertidos*, mientras que los *índices directos*, que son los que vamos a ver ahora, se utilizan como elementos de gestión y control internos.

En un índice directo tenemos la lista de documentos (o de registros) en un orden cronológico (el documento más antiguo primero, por ejemplo) o numérico (del documento número 1 hasta el último). La siguiente figura ilustra esta clase de índices:

Núm. Documento	Contenido
00001	Título: Manual de lenguajes documentales Autor: Blanca Gil
00002	Título: Diseño y construcción de sitios con Dreamweaver 8 Autor: Marc Campbell
...	...
34567	Título: Aprender a razonar Autor: Fina Pizarro

Tabla 1: Representación de un índice directo

Como podemos ver, con este índice, para saber si hay un documento con las palabras "aprender" y "razonar", tendríamos que recorrer decenas de miles de entradas del índice (concretamente 34.567). Lo peor es que, si el índice completo tuviera, por ejemplo, cien mil entradas, habría que recorrer las cien mil entradas del índice para saber si hay más de un documento que cumpla la condición anterior.

Como es fácil suponer, esta clase de índices no mejora mucho en relación al supuesto rastreo en tiempo real de la Web con el que especulábamos antes. Por tanto, un MDB necesita complementar este índice con un índice invertido, que es el que se utiliza realmente para responder a las consultas (mientras que el índice directo se utiliza para aspectos de gestión y administración internos).

5.1.2 Índices invertidos

La estructura de un índice invertido es exactamente la inversa de la anterior (de aquí el nombre), a saber: consiste en una lista ordenada de todas y cada una de las palabras que aparecen en los distintos documentos asociadas a los documentos concretos en los que aparecen. La estructura típica de un índice invertido, por tanto es la siguiente:

Término único	Frecuencia	Ubicación
Aprender	110	(34567, 01, 01) ...
...
Blanca	233	(00001, 02, 01) ...
...
Campbell	6	(00002, 02, 02) ...
...
Razonar	214	(34567, 01, 03) ...
...

Tabla 2: Representación de un fichero invertido

La explicación:

- En **Término único** aparecen las distintas palabras de los documentos, pero solamente aparece una fila por cada palabra (aunque en el conjunto de los documentos aparezca muchas veces).
- En la columna **Frecuencia** tenemos el número total de veces que aparece cada término.
- Por último, en **Ubicación** tenemos una clave en forma de vector donde aparece el número de documento, la zona o campo donde aparece la palabra y el orden de la palabra. Habrá un vector por cada ocurrencia.

Veamos esto último con más detalle. Tomemos el término "Aprender". Tenemos que la columna de **Frecuencia** señala: "110" . Esto significa que "Aprender" aparece 110 veces en el conjunto de los documentos, y que habrá por tanto 110 vectores distintos en la columna **Ubicación** . Nosotros solamente hemos representado uno de ellos para simplificar. Tomemos ahora

este mismo vector [(34567, 01, 01)] y veámoslo con detalle. El primer número, 34567, es el que identifica el documento, el segundo, 01, identifica el campo o zona del documento donde aparece la palabra, en este caso, **Título** .

Otros índices auxiliares ayudan al MDB a recorrer la lista de términos únicos con un pequeño número de operaciones de comparación. Lo mejor de todo es que esta clase de índices hace que el tiempo de respuesta sea virtualmente independiente del número de términos que aparecen en la lista.

El problema principal con estos índices es la cantidad de recursos de hardware que requieren, particularmente, si se debe dar servicio a miles de usuarios simultáneamente.

Actividad 3: Google utiliza alguna versión del tipo de índice invertido que hemos mostrado. Para ver que estos índices son eficaces, pero no permiten hacer milagros, haga una consulta en Google por un término muy frecuente (p.e. Televisión) y compruebe si le resulta posible llegar al resultado número mil (es decir, compruebe si puede llegar a la página de resultados número 100). Probablemente, no podrá. Hasta hora, el motivo no estaba documentado, pero era fácil deducir que Google tiene una cantidad máxima de memoria RAM que puede asignar a cada usuario y por tanto no puede cargar toda la información de sus índices. Una solución técnica consiste en no cargar más de n referencias en memoria para cada consulta.

Para que el MDB pueda construir su índice necesita a su vez un procedimiento que le permita analizar las páginas web y extraer los términos de las mismas que formarán parte del índice, y para ello a su vez necesita poder determinar qué es y qué no es un término.

El algoritmo más fácil determinaría que son términos todas las cadenas de caracteres situadas entre espacios en blanco. Sin embargo este algoritmo produciría muchos errores. Por ejemplo, no seleccionaría las palabras iniciales de cada frase ni sabría tratar las que estuvieran rodeadas de signos ortográficos; por otro lado, tampoco conviene identificar como términos las palabras sin significado propio, como los pronombres y los artículos, etc.

6 Interfaz de consulta

Las funciones que desempeña son las siguientes:

- En primer lugar, el formulario de consulta sirve para enviar la pregunta a otra parte del sistema que compara los términos de la pregunta con el índice invertido y filtra de este modo las páginas web que contienen los términos de la pregunta.
- En segundo lugar, el MDB debe presentar los resultados de la forma más clara y eficiente posible a través de la *Página de resultados del motor de búsqueda (search engine results page)*.

- Por último, pero posiblemente lo más importante de todo, los resultados deben mostrarse en algún orden significativo, y de entrada podemos descartar el orden alfabético o el cronológico, dada su escasa utilidad en un contexto tan heterogéneo y con la inmensa cantidad de documentos que hay en Internet. Aquí interviene el denominado algoritmo de cálculo de relevancia, al que dedicaremos el resto de este apartado.

Actividad 4: Vaya a Yahoo Search (<http://es.search.yahoo.com/>) y acceda al formulario de búsqueda avanzada y examine la opción **Formato**. ¿Cuántos tipos distintos de documentos permite buscar? (En el momento de preparar esta unidad podían buscarse siete formatos distintos). Haga pruebas con la misma palabra clave pero con distintos formatos. Observe el número total de formatos disponibles y si probando con distintos formatos ha obtenido mejor (o peor) relevancia.

6.1 Cálculo de relevancia

En recuperación de información se considera que los distintos documentos de un fondo presentan un grado de relevancia diferente para cada pregunta o necesidad de información. Ante una pregunta dada, en un extremo tendremos a los documentos cuya relevancia será cero (por no tener ninguna relación con la pregunta). En el otro extremos tendremos (con suerte) documentos cuya relevancia será total (o casi total). En zonas intermedias tendremos documentos con diversos grados de relevancia. Si la expresamos con tanto por ciento, tendremos un espectro que irá del 0% de relevancia al 90% o al 100% de relevancia con zonas intermedias donde habrá documentos con el 10%, el 25% o el 70% de relevancia, etc.

Lo más lógico será presentar la página de resultados con los documentos ordenados según su grado de relevancia. Esta ordenación se denomina también *ranking*. De este modo, aunque haya miles y miles de documentos relevantes en la respuesta, podremos limitarnos a consultar los 10 o 20 primeros, puesto que se supone que son los más relevantes. Otros modos de ordenación, como el alfabético o el cronológico no nos permitirían tal economía de consulta. Para estar seguros que accedemos a los más relevantes itendríamos que examinar la lista completa!

Los MDB actuales suelen combinar dos grupos de criterios para determinar la relevancia de una página web:

- Criterios internos o intrínsecos
- Criterios externos o de popularidad

6.1.1 Criterios intrínsecos

Los criterios internos se refieren principalmente a aspectos estadísticos o de frecuencia de ocurrencia de la palabra (o palabras) clave de la pregunta. En concreto, las páginas web que tengan mayor densidad de la palabra clave, serán más relevantes.

Otros aspectos, como que la palabra clave aparezca rodeada de etiquetas como <h1> o <h2> también otorga mayor importancia relativa a la página correspondiente, y sobre todo, al hecho de que la palabra clave forme parte de la URL de la página.

6.1.2 Criterios externos

Los criterios externos se refieren al resultado que pueda arrojar el análisis de los enlaces de entrada de la página considerada.

Cada MDB tiene sus propios conjuntos de criterios y sus propias reglas para asignar pesos a cada criterio, pero en general, el número de enlaces que recibe una página suele ser uno de los más importantes, al menos a igualdad de los otros factores.

7 Página de resultados

Por último, la respuesta del motor con la lista de documentos organizados en un ranking debe presentarse de alguna forma. De este aspecto se ocupa precisamente la página de resultados (PDR a partir de ahora).

Las funciones de la PDR, más allá de poder hacer clic en ellos y, por alguna razón, suelen pasar desapercibidas. Tal vez está bien que sea así, porque cada vez que se han hecho intentos en contra, parece el motor ha perdido popularidad.

En este sentido, a simple vista, la PDR ofrece una lista de páginas o documentos y ya está. Sin embargo, una mirada atenta a la PDR de, por ejemplo, Google, nos muestra las siguientes informaciones y opciones:

- (1) Número total de páginas que contienen la palabra clave.
- (2) La sección de resultados de pago (resultados patrocinados).
- (3) La sección de resultados no pagados (resultados orgánicos). De cada uno de estos resultados, tenemos la siguiente información:
 - Título de la página (o del documento)
 - El tipo del documento (cuando no es HTML)

- Unas líneas de descripción del contenido del documento
- URL de la página
- Tamaño de la página web (no de los archivos enlazados)

(4) Y las siguientes opciones:

- Obtener una versión traducida de la página con traducción automática (o sea, en general muy deficiente)
- Ver la página en la caché de Google
- Buscar páginas con contenidos similares
- Navegación secuencial entre los resultados o yendo a una página de resultados concreta (hasta la página 90 más o menos)
- Restringir la siguiente búsqueda a los resultados obtenidos

Actividad 5: Vaya a Ask (www.ask.com), haga una pregunta por cualquier término (p.e. "television digital") y observe la página de resultados. Compárela con la página de resultados de Google (vea la lista superior). ¿Ofrece las mismas opciones? ¿Hay alguna opción en Ask que no esté en Google -o al revés-? ¿Cuál le parece mejor?

8 Conclusiones

Conocer los aspectos estructurales y funcionales que hemos examinado aquí nos ayudará sin ninguna duda en nuestra relación con los motores de búsqueda. Conocer un poco mejor los fundamentos de los motores nos ayudará a llevar a cabo mejores búsquedas, pero seguramente, nos ayudará también a desarrollar sitios web más amistosos con los motores de búsqueda, y por supuesto, debería ayudarnos a desarrollar mejores campañas de posicionamiento.

9 Bibliografía

Abadal, E.; Codina, L. "Recuperación de información". En: *Bases de datos documentales: Características, funciones y método*. Madrid: Síntesis, 2005

Chu, H. *Information representation and retrieval in the Digital Age*. Medford: Information Today, 2003

Codina, L. "Modelo conceptual de un sistema de información documental". *Revista española de documentación científica* , v. 17, n. 4, Octubre-diciembre 1994, p. 44-53

Codina, L. "Cómo funcionan los servicios de búsqueda en Internet: un informe especial para navegantes y creadores de información". Partes I y II. *Information world en español* , 1997

Kowalski, G. "Automatic indexing". En: *Information Retrieval Systems: Theory and Implementation* . Boston: Kluwer, 1997

Liu, B. "Information retrieval and Web Search. En: *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data* . Berlin: Springer, 2007

Milstein, S.; Biersdorfer, J:D.; MacDonald, M. *Google: The Missign Manual*. Sebastopol: O`Reilly, 2006

Título	Artículo P235. El ciclo de vida de una campaña de posicionamiento
Autor	Lluís Codina
Palabras clave	Posicionamiento web, ciclo de vida, palabras clave, altas en motores, contenidos
Descripción	El posicionamiento en buscadores es un proceso dinámico que presenta una serie de fases. Al mismo tiempo, no existe en realidad un final en el posicionamiento de una web ya que es necesario monitorizar con cierta frecuencia el éxito de las acciones de posicionamiento y realizar acciones de corrección si es necesario.
Objetivos	El principal objetivo de la unidad didáctica es presentar las distintas fases de las que se compone una campaña de posicionamiento web, así como las diferentes herramientas y procesos que acompañan a cada fase.

Sumario

1	Introducción.....	2
2	Las fases de la campaña.....	2
3	Fase de análisis.....	2
3.1	Palabras Clave.....	2
3.2	Fase de diseño.....	5
3.3	Fase de implantación.....	6
4	Fase de seguimiento.....	7
4.1	Calendario.....	7
4.2	Herramientas.....	7
5	Conclusiones.....	9
6	Bibliografía.....	9

1 Introducción

El posicionamiento en buscadores es un proceso dinámico. Tiene unas fases determinadas que denominamos "ciclo de vida" pero, no existe en realidad un final en el posicionamiento de una web, ya que cambios en la propia web, en los algoritmos de los buscadores, etc., hacen necesario mantener una política de monitorización y actualización de las campañas de posicionamiento de una web.

2 Las fases de la campaña

Entendemos que estas fases siguen la fórmula "3+1" porque incluyen tres fases de desarrollo más una fase de seguimiento:

- 1. Análisis** del sitio
- 2. Diseño** de las estrategias
- 3. Implementación** de las medidas adoptadas

Pasado un tiempo (al menos un mes) de la aplicación de las fases anteriores deberá aplicarse una cuarta fase:

- 4. Seguimiento** de los resultados

Es necesario destacar que deberá aplicarse el ciclo de vida de manera individualizada para tantas secciones de interés estratégico como tenga el sitio web. La razón es la siguiente: dada la naturaleza de la Web, no existe ninguna razón por la cual debamos suponer que los visitantes de un sitio accederán todos a través de la página principal del mismo. En segundo lugar, si un sitio web es rico en contenidos, se supone que estaremos interesados en promover las distintas secciones de modo independiente y no solamente la página principal.

3 Fase de análisis

3.1 Palabras Clave

El objetivo más importante de esta fase es determinar las palabras clave para las cuales se desea optimizar el sitio, así como identificar a los posibles sitios que pueden enlazar de *forma justificada* a nuestro sitio.

Recordemos que denominamos palabras clave a los términos de consulta que utilizan los usuarios de los motores de búsqueda. Por el mismo motivo, se denomina palabra clave a los términos presentes en el contenido de un sitio que, eventualmente, pueden coincidir con las palabras usadas por los usuarios cuando realizan sus búsquedas. Por supuesto, las palabras clave son también los términos que figuran en los índices de los motores de búsqueda.

Una palabra clave puede ser una palabra simple (p.e. "Arte") o una expresión compuesta (p.e. "Historia del Arte"). Puede ser un nombre común ("museo") o un nombre propio ("Louvre"). Puede expresar conceptos ("belleza") o cosas tangibles ("casas"). También pueden corresponder a ramas del conocimiento ("arquitectura") y actividades humanas ("reparación de ordenadores").

Las palabras clave serán elegidas en función de los tres elementos principales que definen a un sitio: el contenido, los objetivos y el público, haciendo énfasis en este último elemento, es decir el público al que está destinada la web objeto de la campaña.

Se recomienda comenzar por buscar tres o cuatro palabras clave principales relacionadas con el sitio. Por ejemplo, en la web de un museo de cinematografía, podrían ser las palabras "museos" y "cine"; y después derivar de ellas cinco o seis palabras clave adicionales más específicas. Por último, deben añadirse nombres propios, p.e., el nombre de la institución y algún topónimo si hace al caso.

En esta tarea es importante involucrar al departamento comercial en caso de que se trate del sitio de una empresa. En concreto, para ayudar a la determinación de las palabras clave existen diversos procedimientos:

- El análisis del propio sitio por un experto humano y/o con programas específicos como IBP (www.axandra.com).
- El análisis de frecuencias de la diversas palabras clave en la web relacionadas con el sitio y de sus diversos sinónimos utilizando los propios motores de búsqueda, algunos de los cuales suelen ofrecer variantes del término buscado (como Ask o Clusty).
- El *benchmarking* por comparación de las palabras clave utilizadas en las páginas web de la mejor competencia.
- El análisis de las preguntas de los usuarios, si nuestro sitio dispone de un motor de búsqueda interno.
- El análisis de las palabras más buscadas y de sus variantes respectivas que podemos obtener a través de las herramientas de algunos motores de búsqueda (en particular, la herramienta para webmasters de Google: <http://www.google.es/webmasters/>)

Actividad 1 : Entre en el sitio para webmasters de Google:

<http://www.google.es/webmasters/>

y realice un recorrido libre por las principales herramientas que ofrece.

Alternativa: descargue el programa IBP (www.axandra.com) y obtenga un informe del sitio de su empresa con el mismo.

Actividad 2: Establezca una lista de cinco palabras clave para el sitio web de su empresa o corporación. *Alternativa:* entre en la web de la Filmoteca Española (<http://www.mcu.es/cine/MC/FE/>) y haga una lista de las cinco palabras clave para las que cree que debería estar optimizada. En cualquiera de los dos casos, haga búsquedas en Google y compruebe si la web de su empresa (o la web de la Filmoteca Española) aparece en la primera página de resultados.

El resultado de los análisis anteriores debería ser no solamente un conjunto de palabras clave, sino también dos frases descriptivas del sitio. Una de las frases, optativa, y más breve, suele denominarse *tagline*. Suele figurar en un lugar destacado de la página principal del sitio y se suele colocar en un lugar cercano al título principal. La *tagline* se puede repetir en todas las páginas del sitio. Esta frase forma parte del contenido visible del sitio. No todos los sitios consideran necesario disponer de una tagline (o al menos, de facto, no disponen de ella).

La segunda frase forma parte del contenido invisible, ya que se usará como contenido del atributo *description* en la etiqueta <meta> del código fuente y se considera imprescindible. Idealmente, la frase *description* en cada página del sitio debería ser similar, pero no idéntica, puesto que cada página se supone que tiene un contenido que es, en parte similar, pero en parte distinto. Puede adoptarse la solución de generar dos frases: una frase corta que se repetirá en toda la web, y otra que se añadirá a continuación y será específica de cada página particular. Podemos preparar una ficha de *identidad* del sitio como ésta:

Ficha de identidad del sitio

Tagline
<meta> Description
<meta> Keywords

Para nuestro hipotético Museo de Cinematografía, la ficha podría quedar así:

Ficha de identidad Museo de la Cinematografía (sitio ficticio)

Tagline	El cine es el arte de nuestra vida
<meta> Description	El Museo de la Cinematografía Luis Buñuel propone un recorrido por la historia del cine español que incluye más de 500 objetos presentados en diversas salas, así como presenta exposiciones temporales sobre conceptos universales del cine.

<meta> Keywords	Museo de Cinematografía, Luis Buñuel, Teruel, Aragón, Cine, Historia del cine, Exposiciones, Films, Cámaras cinematográficas, Fotografía, Historia de la imagen, Imágenes en movimiento, Archivo cinematográfico, Cine sonoro, Cine mudo, Educación, Medios de comunicación, Comunicación Audiovisual
-----------------	---

La ficha anterior debe ser aprobada por la dirección del organismo que publica el sitio y, por supuesto, si se trata de una empresa, debe participar en su redacción el departamento comercial.

Actividad 3: Haga la ficha de identidad del sitio de su empresa o de la web de la Filmoteca Española (<http://www.mcu.es/cine/MC/FE/>). *Alternativa:* puede hacer la ficha de cualquier otro sitio que sea de su interés, si lo prefiere.

El segundo producto de esta fase será una lista de directorios o guías temáticas, etc. sobre cinematografías o aspectos relacionados que contengan una sección de enlaces y que, por tanto, podrían enlazar al sitio.

3.2 Fase de diseño

El objetivo es presentar propuestas de *optimización* específicas sobre los siguientes aspectos:

Código fuente del sitio:

- Palabras clave en el atributo *keyword* y descripciones en atributo *description* dentro de los distintos elementos < meta>, así como contenido del elemento <title> en la sección <head> de la página principal (y de las demás páginas de la organización), así como el atributo *author* si se considera oportuno especificar el nombre o nombres de los autores personales o corporativos del sitio.
- Aparición de las palabras clave más importantes en el atributo *title* en los enlaces de navegación interna y en los enlaces externos y en el atributo *alt* en los elementos y *title* en <object>
- Aparición de las palabras clave principales en títulos de los enlaces y en títulos de secciones (etiquetas <h1>)

Contenido del sitio:

- Ubicación de la *tagline* en la página principal y, eventualmente, en las demás páginas del sitio.
- Preparación de una o más frases o párrafos que deberán contener las principales palabras clave del sitio en la página principal del sitio.

- Planificación de la creación regular de contenidos de calidad para publicar en el sitio y que estén vinculados con las palabras clave del sitio y donde éstas ocurran con naturalidad. Por ejemplo, planificar que, una vez al mes (o cada semana) se publicará un informe sobre las distintas colecciones presentes en el museo, que puede incluir imágenes. Organizar una ciclo de conferencias anual a celebrar en el museo y publicar los contenidos de las conferencias, etc.

Enlaces:

- Propuesta de un apartado de enlaces hacia sitios similares (nunca en la primera página del sitio ni tan solo en el primer nivel)
- Lista de sitios web que podrían enlazar nuestro sitio con el correo electrónico del administrador del sitio

Actividad 4 : Revise el código fuente de la página principal de su empresa y compruebe la adecuación del mismo en relación a los metados: etiqueta <title> y etiquetas <meta>. Alternativa: haga este análisis con la página principal de la web de la Filmoteca Española (<http://www.mcu.es/cine/MC/FE/>).

3.3 Fase de implantación

Subfase I: Aprovisionamiento de enlaces de entrada

Antes de dar de alta el sitio en los motores que previsiblemente nos van a aportar la mayoría del tráfico, es importante proveerse de enlaces de entrada. De este modo, cuando los motores entren en nuestro sitio y lo analicen, descubrirán que ya tenemos ese ingrediente tan importante para obtener buenos rankings en las páginas de resultados. Podemos aplicar estas acciones:

- Altas en los directorios generalistas: Yahoo, Dmoz
- Altas en directorios y portales temáticos o regionales. Un buen punto de partida si no conocemos ning uno puede consistir en consultar el directorio Dmoz.
- Captación de enlaces corporativos y de sitios de temática similar. Además de los enlaces que nuestro sitio puede obtener de los directorios anteriores, sin duda podremos obtener enlaces de otros sitios; p.e. de otras webs vinculadas con nuestra institución, así como de otras webs con contenidos similares (siempre es conveniente demostrar que nosotros ya los hemos enlazado).

Subfase II: Altas en motores

- Altas en, al menos, los siguientes motores generalistas: Google, Yahoo, Live y Ask

En toda esta fase es aconsejable utilizar dos juegos de fichas:

- La ficha de identidad del sitio (para entrar siempre las mismas descripciones en los distintos motores)
- Ficha de control de altas, a fin de saber en todo momento en qué motores o directorios se han dado las altas y cuándo se hizo.

Por último, creemos conveniente que el orden seguido en las actividades anteriores sea el mismo que se ha presentado aquí, a saber: primero conseguir enlaces y después dar el alta en los motores, siempre que entre las dos fases se dejen transcurrir dos semanas (tampoco conviene esperar más).

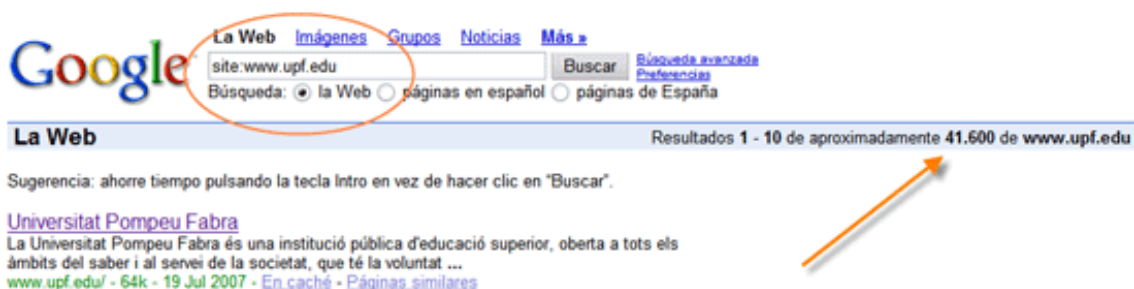
4 Fase de seguimiento

4.1 Calendario

Entre el final de la campaña y el primer seguimiento es aconsejable esperar al menos un mes. Monitorizar resultados antes puede resultar a la vez frustrante y engañoso. Los motores de búsqueda pueden tardar semanas en indizar el sitio desde que recibieron el alta. Para seguimientos posteriores muchos profesionales del posicionamiento incluso aconsejan períodos más amplios, típicamente, aconsejan hacer un seguimiento cada tres meses.

4.2 Herramientas

Lo primero que haremos será constatar que nuestro sitio ha sido indizado. Para ello, los motores disponen de un comando que permite conocer si nuestro sitio figura en su índice y cuantas páginas de nuestro sitio han sido indizadas. Tanto en Google como en Yahoo el comando es site. Por ejemplo, si queremos saber si la web de la UPF está indizada en Google y cuántas páginas ha indizado podemos utilizar este comando del modo que muestra la captura:



Actividad 5 : Haga el análisis del número de páginas que indiza Google del dominio de su empresa utilizando el comando `site` . Alternativa: puede realizar este análisis del dominio de la Filmoteca Española (`site:www.mcu.es/cine/MC/FE/`).

Una vez sepamos que nuestro sitio ha sido indizado, la comprobación más importante consistirá en hacer búsquedas por las diversas palabras clave para las cuales hemos optimizado nuestro sitio y anotar en qué posición queda en la página de resultados de cada motor. El objetivo debe ser aparecer en las primeras páginas de resultados y, a ser posible, en la primera. Si no lo hemos conseguido, las causas pueden ser de tres tipos:

- **Técnicas:** tal vez el sitio presenta algún problema de indización a los motores. Por ejemplo, el sitio puede usar tecnologías obsoletas, como *frames* (marcos). La prueba anterior (el comando `site`) pondrá esto de manifiesto si, por ejemplo, solamente una parte muy pequeña de la web ha sido indizada.
- **De contenido:** Tal vez el sitio presenta escaso contenido (con ocurrencia de las palabras clave).
- **De popularidad:** Tal vez el sitio dispone de pocos enlaces de entrada de sitios que a su vez sean muy populares. Aquí podemos utilizar una medida de popularidad muy significativa, el *PageRank* de Google que puede conocerse instalando la *toolbar* de Google (toolbar.google.com). Igualmente, podemos monitorizar el *TraffickRank* de nuestro sitio a través de la *toolbar* de Alexa (www.alexa.com).

Las causas técnicas deberían poder resolverse sin mayores problemas. El uso de *frames* para crear páginas está de hecho, desaconsejado desde hace tiempo, y los modernos sistemas de creación de páginas web presentan muchas opciones para ello, la principal, el uso de hojas de estilo. Los sistemas de gestión de contenidos también se pueden parametrizar para que creen menús con enlaces fáciles de seguir por los motores de búsqueda, etc.

En cuanto a los problemas de contenido, no existe una solución fácil ni rápida (ni económica, debemos añadir), y aunque a muchos les pueda extrañar, es justo que sea así. Nadie espera en realidad que los motores de búsqueda concedan las primera posiciones a webs con contenidos irrelevantes para la pregunta. Por tanto, si nuestra web posee pocos contenidos sobre las palabras clave por los que deseamos obtener buenas posiciones, la única solución es planificar un calendario de producción y publicación de tales contenidos. Ni siquiera sirve añadir toneladas de metadatos (en etiquetas `<meta>`) porque los motores de búsqueda han aprendido a desconfiar de ellos. Para los motores de búsqueda es una cuestión de supervivencia promocionar a buenas posiciones solamente a sitios con buenos contenidos en calidad y en cantidad.

Por lo que hace a la popularidad, si se han dado los pasos del punto 3.3 para obtener enlaces y pese a todo nuestro sitio dispone de pocos enlaces de entrada, la razón principal solamente suele ser simplemente que los otros webmasters no han considerado interesante nuestro sitio, bien por tener pocos contenidos, bien por ser de poca calidad o poco originales o por una combinación de todo esto. Por tanto, el problema de la popularidad como se puede ver va ligado con el de los contenidos y, en teoría, resolver el primero debería resolver el segundo, siempre que se den los pasos indicados en el punto 3.3.

5 Conclusiones

A modo de conclusiones, podemos señalar los siguientes puntos:

1. Las campañas de posicionamiento tienen su inicio en la selección de palabras y frases clave para las cuales deseamos optimizar nuestro sitio de cara a los motores de búsqueda.
2. Nuestro sitio web debe tener contenidos de texto: no sirven los gráficos o las animaciones en flash.
3. Nuestro sitio debe ser optimizado tanto a nivel de contenidos como de código fuente para las palabras clave seleccionadas.
4. La obtención de enlaces de entrada puede ser, en algunos casos, más del cincuenta por ciento del éxito en una campaña de posicionamiento.
5. Los contenidos, en la Web son siempre el elemento clave. Las actualizaciones, es decir, la provisión frecuente de nuevos contenidos refuerza este papel clave.
6. Los buenos contenidos, sin embargo, no sirven de nada si nuestra web presenta problemas técnicos que impiden el acceso a los motores de búsqueda.
7. Igualmente, los buenos contenidos no sirven de nada, al menos a corto y medio plazo, si no adoptamos una política activa de posicionamiento como la presentada aquí.

6 Bibliografía

Codina, L. "El posicionamiento web y su importancia para las empresas". *Indicador de economía*, Marzo 2007. (Acceso: www.lluiscodina.com/posicionamiento2007v2.pdf)

Grappone, J.; Couzin, G. "Your one-month prep: Baseline and Keywords". *Search engine optimization: an hour a day*. Indianapolis: Sibex, 2006.

Kent, P. "Part II: Building search-engine-friendly sites". *Search engine optimization for dummies*. Indianapolis: Wiley, 2006.

Sweeney, S. "Designin your site to be search engine friendly". *101 Ways to promote your web site*. Gulf Breeze: Maxpress, 2005.

Sweeney, S. "Search engine and directory submissions". *101 Ways to promote your web site*. Gulf Breeze: Maxpress, 2005.

Título	Artículo M519. Fundamentos del márketing en buscadores
Autor	Cristòfol Rovira
Palabras clave	Márketing en Internet, márketing en buscadores, pago por clic, resultados patrocinados, palabras clave, publicidad en buscadores
Descripción	Se analizan las principales características de los servicios de publicidad en buscadores haciendo especial incidencia en Adwords de Google y en Yahoo Sponsored Search
Objetivos	El objetivo de esta unidad es conocer los principales conceptos relacionados con la temática del módulo como "Márketing en Internet" y "Márketing en buscadores". También se analizarán los distintos modelos de facturación de los enlaces patrocinados y se estudiará con detenimiento como es la estructura y el funcionamiento de una cuenta para gestionar este tipo de servicios.

Sumario

1 Introducción.....	2
2 Márketing en Internet.....	2
3 Modelos y formas de pago.....	4
4 Resultados orgánicos, resultados patrocinados.....	5
5 El mercado de márketing en buscadores.....	6
6 Funcionamiento de una cuenta de publicidad en buscadores.....	7
6.1 Estructura.....	7
6.2 Pago por clic.....	8
6.3 Palabras clave.....	10
7 Conclusiones.....	11
8 Bibliografía.....	12

1 Introducción

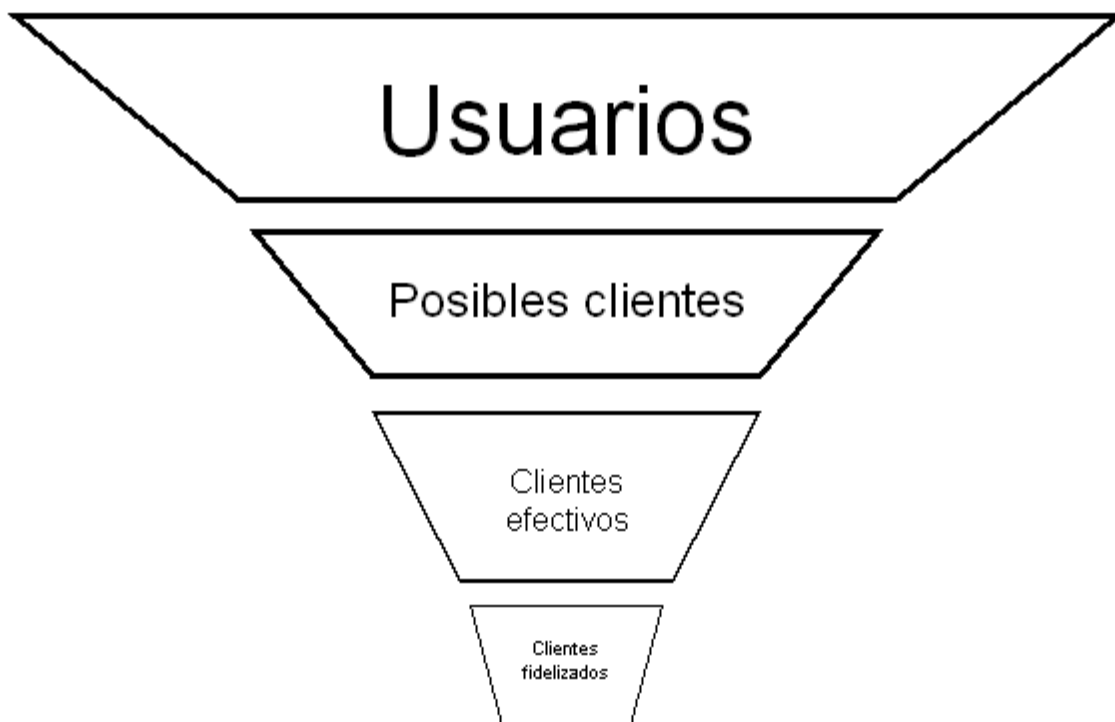
El objetivo de esta unidad es conocer los principales conceptos relacionados con la temática del módulo como “Márqueting en Internet” y “Márqueting en buscadores”. También se analizarán los distintos modelos de facturación de los enlaces patrocinados y se estudiará con detenimiento como es la estructura y el funcionamiento de una cuenta para gestionar este tipo de servicios.

2 Márqueting en Internet

Para que una empresa funcione en Internet es necesario que haga bien los siguientes cuatro procesos básicos:

- Captación de usuarios hacia su sede web
- Conversión de los usuarios en posibles clientes (leads)
- Conversión de los potenciales clientes en clientes efectivos
- Fidelización de los clientes

Cada uno de estos procesos depende del anterior y al mismo tiempo el número de personas implicadas va disminuyendo, de forma que aproximadamente solo un 1% de los usuarios pasan a clientes, tal como podemos ver en el siguiente gráfico:



Un punto crítico de todo el proceso es la obtención de los usuarios que iniciaran la cadena. En este punto se sitúan tanto el márketing en buscadores (SEM Search Engine Marketing) como el posicionamiento en buscadores (SEO Search Engine Optimization).

El posicionamiento en buscadores consiste en el conjunto de procesos para conseguir mejorar la posición de una sede web en el listado de resultados ante unas determinadas palabras clave y en un determinado buscador. El listado de resultados depende de multitud de factores, entre los más importantes está el número de enlaces que recibe la sede web. Se denominan resultados orgánicos o naturales a estos listados generados al aplicar los criterios objetivos que el buscador considera adecuados para obtener en los primeros lugares a las webs de mayor calidad y al mismo tiempo las más adecuadas a la búsqueda realizada. El posicionamiento en buscadores no implica en ningún momento pagar al buscador para conseguir un mejor posición.

Por otro lado la publicidad en buscadores incluye los servicios y procesos que tienen como objetivo la gestión de anuncios en los buscadores de una determinada sede web. Estos anuncios, llamados enlaces patrocinados, están situados en el listado de resultados de los buscadores, pero en una zona claramente diferenciada de los resultados orgánicos. La publicidad en buscadores es de pago y se factura por cada clic que los usuarios del buscador hagan en el enlace patrocinado que conduce a la sede web patrocinada.

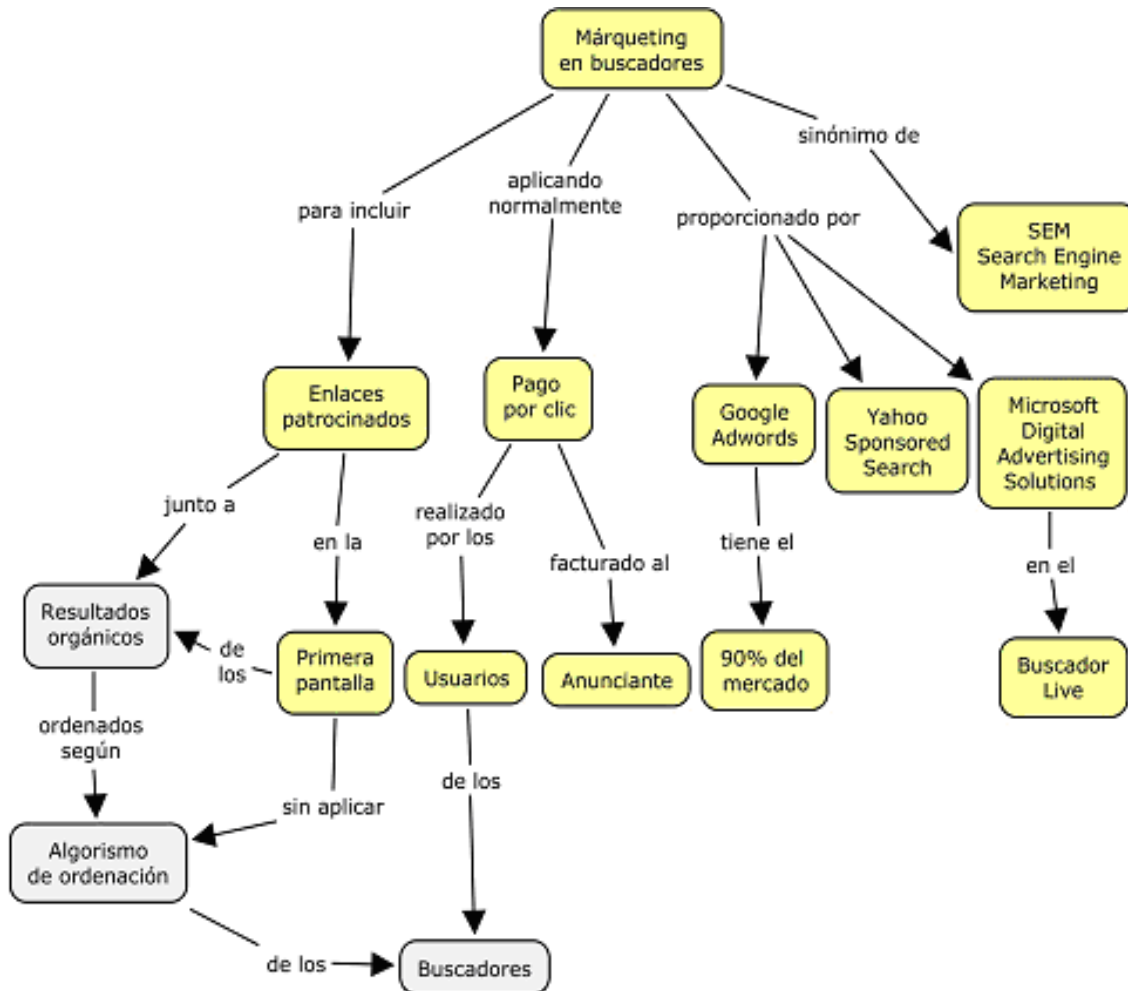
Existen diversos servicios de publicidad en buscadores asociados a los principales buscadores disponibles en la red:

- Google con el servicio Adwords (tiene el 90% del mercado)
- Yahoo con el servicio Sponsored Search
- Live con el servicio Microsoft Digital Advertising Solutions

El márketing en buscadores será el eje central de esta unidad didáctica, veremos un par de definiciones y analizaremos sus características, funciones y objetivos. El márketing en buscadores forma parte de la joven disciplina del márketing online, márketing digital o márketing en Internet. El márketing en Internet incluye, entre otros, los siguientes métodos o técnicas publicitarias:

- El marketing en buscadores
- El posicionamiento en buscadores
- La publicidad de pago por clic
- Los banners o avisos en páginas web
- Los envíos de correo masivos
- El márketing de bitácoras (blogs)

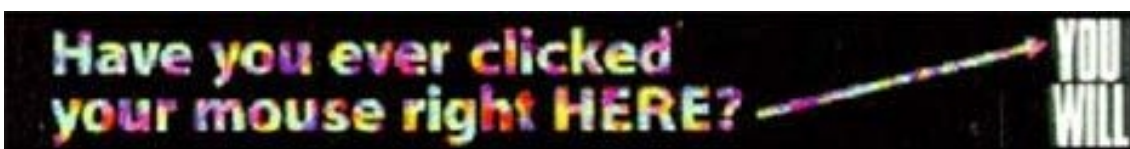
De entre todas estas técnicas, la publicidad de pago por clic en el contexto de los buscadores es la más efectiva, la que se dedica más presupuesto y la que tiene mayores expectativas de crecimiento.



3 Modelos y formas de pago

Hay dos referencias de la joven historia del Márketing en Internet especialmente significativas:

- 25 de octubre de 1994: se publica el primer banner publicitario en la cabecera de la web HotWired.com. Es la fecha del nacimiento del Márketing en Internet.
- A principios del 2000 Google lanza el servicio Adwords, revolucionando el Márketing en Internet por sus prestaciones, en especial por el control que pueden ejercer los usuarios d el gasto del presupuesto invertido en este tipo de publicidad.



Primer banner publicitario, 1994

Entre 1994 y el 2000 se ensayaron distintos modelos de publicidad en Internet, algunos de ellos muy agresivos, como las ventanas superpuestas automáticas (pop-up). Desde la perspectiva de las empresas, los modelos de publicidad en la red son analizados en función de su coste y de la forma en que se factura el servicio. Hay dos formas básicas de facturación:

- Por número de "impresiones" (Cost Per Impression CPI), o sea por el número de veces que un mensaje publicitario es mostrado en una página web para que un usuario lo vea.
- Por el número de clics (Cost Per Click CPC) que realiza un usuario en el mensaje publicitario y que implicará una visita a la web publicitada y por tanto una consulta al servicio o producto que contiene.

El pago por impresiones favorece a la web que proporciona el soporte para la publicidad puesto que obtendrá ingresos con independencia del comportamiento del usuario y en función del número de vistas. En cambio, el pago por clic favorece a la empresa que se publicita puesto que sólo pagará si efectivamente consigue una vista.

Si el anuncio es coherente con el contenido de la web que soporta la publicidad, entonces los usuarios probablemente la considerarán atractivo y por tanto el porcentaje de visitantes que harán clic será alto. Pero este dato muchas veces no es posible obtenerlo a priori y por tanto, para la empresa que se publicita siempre es mucho mejor el pago por clic porque asume un riesgo mínimo.

El Márqueting en Buscadores es la expresión máxima del encaje entre el contenido de la web que soporta el anuncio y el mensaje del anuncio ya que un buscador tiene un contenido neutro que el propio usuario define con sus búsquedas y son precisamente estas búsquedas, y las palabras clave usadas, las que disparan los anuncios.

Actividad 1: Supongamos una sede web con 5000 visitas diarias. Si ponemos un anuncio en esta sede web y nos cobran 0,0001 euros por impresión ¿qué porcentaje de visitantes deberían hacer clic en el anuncio para que me costase lo mismo pagar 0,02 euros por clic? A la vista del resultado, ¿en qué casos será mejor aceptar una campaña de pago por número de impresiones?

4 Resultados orgánicos, resultados patrocinados

En los listados de resultados de los buscadores aparece claramente diferenciadas dos zonas, la primera la zona de resultados obtenidos en función de las palabras clave introducidas en la búsqueda y ordenados de acuerdo con el algoritmo de ordenación correspondiente a cada buscador. Como se ha visto, este algoritmo depende de muchos factores, pero ninguno de ellos es el pago realizado por las empresas que aparecen en el listado. Son los resultados orgánicos.

En el listado de resultados aparecen también el listado de anuncios bajo el título de “enlaces patrocinados”. Están situados en una zona bien diferenciada y también son activados en función de las palabras clave usadas.

Google y Yahoo tienen especial interés en diferenciar claramente las dos zonas de resultados y aunque se indique de forma clara que se trata de enlaces patrocinados, muchos usuarios no identifican que se trate de publicidad. La zona de enlaces patrocinados en ambos buscadores es idéntica en forma de L invertida y con la frase “Enlaces patrocinados” dos veces, una para la banda superior y otra para la banda derecha.

Cuando la búsqueda realizada se usan palabras muy comerciales con una gran competencia, pueden aparecer hasta 10 enlaces en la primera y sucesivas páginas del listado de resultados y un número indeterminado en páginas adicionales accesibles mediante un enlace titulado “Más enlaces patrocinados”. Como veremos en las siguientes unidades, es muy importante que un enlace patrocinado salga en las primeras posiciones del listado ya que el número de clics disminuye de forma muy rápida en función de esta posición.

Actividad 2: Haga una búsqueda en Google en español y otra en Yahoo en español con palabras con gran competencia y analice como gestionan ambos buscadores la presencia de muchos anuncios en una búsqueda. ¿Las dos zonas de anuncios son idénticas? ¿existen páginas adicionales con anuncios a parte de la página de resultados del buscador?

5 El mercado de márketing en buscadores

El mercado mundial de tráfico en los buscadores se lo reparten Google, Yahoo y Live con más del 90% de las búsquedas realizadas. Según los últimos datos de enero del 2008, Google tendría una cuota de mercado entre el 60% y el 65%, Yahoo el 20% y Live un 10%.

Estos porcentajes se reflejan también en la cuota de mercado en publicidad en buscadores que también está a favor de Google pero con mayor diferencia. Debido a que Google Adwords fue el primer servicio de publicidad en buscadores con altas prestaciones, los datos que se manejan actualmente son que el 90% del mercado está en manos de Google.

En el último año Yahoo ha hecho un gran esfuerzo para ofrecer un servicio de prestaciones similares a Google incorporando muchas mejoras imitando en muchos detalles la cuenta de Google. No obstante, Google es el líder con diferencia de este mercado, especialmente en España.

6 Funcionamiento de una cuenta de publicidad en buscadores

6.1 Estructura

Una cuenta de Márqueting en buscadores tienen una estructura muy bien definida que aplican tanto Google Adwords como Yahoo Sponsored Search formada por:

- Un conjunto de campañas publicitarias, cada una de ellas relacionada con un producto, servicio o marca comercial
- Cada campaña está formada por diversos grupos de anuncios con el objetivo de tener un grupo de anuncios para cada aspecto o elemento concreto del producto o servicio que queremos promocionar
- Cada grupo de anuncios está formada por diversas variaciones de un mismo anuncio referido a un mismo aspecto concreto.
- Finalmente cada anuncio tiene asociadas un conjunto de palabras de clave que serán las encargadas de "disparar" su impresión cuando un usuario las utilice

Un ejemplo sencillo remitirá ver más claramente cada elemento. Supongamos una empresa pequeña que inicia su andadura en la promoción en buscadores y que da de alta una campaña sencilla para un solo producto. La estructura de la cuenta para esta empresa será la siguiente:

- Una sola campaña sobre el producto a promocionar
- Un solo grupo de anuncios formado por un solo anuncio
- 10 palabras clave asociadas a este anuncio

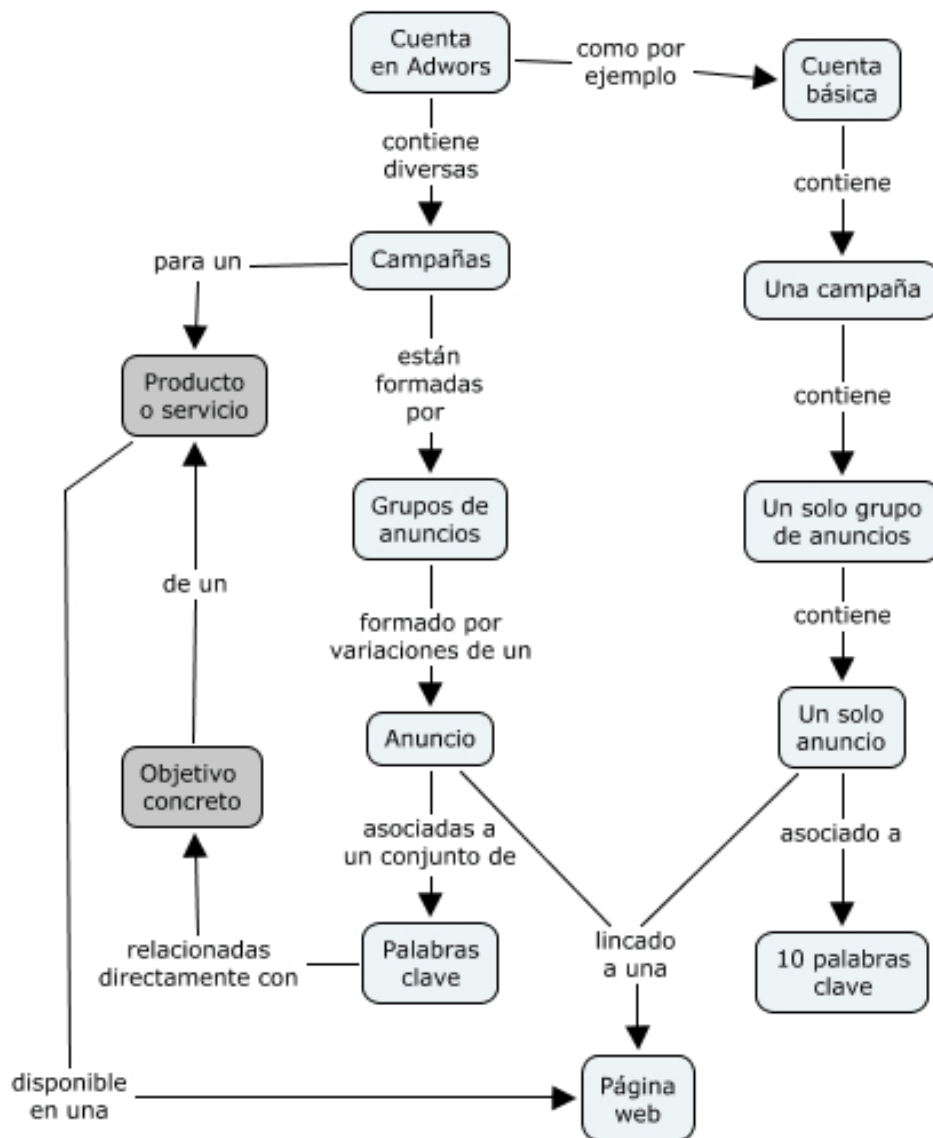
Cuando los responsables del márqueting en buscadores quieran optimizar esta campaña es posible que en un primer momento decidan mejorar la estructura inicial con los siguientes cambios:

- Una sola campaña sobre el producto a promocionar
- Un solo grupo de anuncios formado por diversas variaciones del mismo anuncio
- Las mismas 10 palabras clave asociadas a todos los anuncios del grupo

Un segundo cambio o ampliación que se podría realizar para seguir optimizando la campaña sería:

- Una sola campaña sobre el producto a promocionar
- Dos grupos de anuncios sobre dos orientaciones distintas de la campaña sobre el mismo producto, la primera acentuando sus propiedades y la segunda con una promoción.
- Diversas variaciones de anuncios en cada grupo
- 10 palabras clave asociadas a todos los anuncios del grupo 1 y 10 palabras clave asociadas a los anuncios del grupo 2

Es interesante resaltar que un anuncio tiene asociadas un grupo de palabras clave, pero si existen diversas variaciones sobre este anuncio, formando un grupo de anuncios, todos los anuncios quedan asociados a las mismas palabras. Por tanto, si queremos poner dos anuncios asociados a palabras clave distintas tendremos que crear un grupo específico para cada anuncio.



6.2 Pago por clic

La publicidad en buscadores se facturan en función del número de clics que realizan los usuarios en los enlaces patrocinados. Esta forma de facturación es muy venajosa para el enunciante puesto que el dinero pagado implica visitas reales en la web promocionada. No obstante, para que haya un clic antes tiene que haber antes una presentación del anuncio en el listado de resultados. A la aparición de un anuncio en el listado de resultados se denomina "impresión del

anuncio" y hay dos factores básicos que regulan esta impresión de anuncios:

- Las palabras clave que el anunciante determina que están asociadas al anuncio
- El dinero que está dispuesto a pagar el anunciante por cada clic de un usuario del buscador en el anuncio

Cada anuncio tiene un listado de palabras clave asociadas que en el momento que sean usadas en una búsqueda dispararán la impresión del anuncio. Por ejemplo, un banco puede tener anuncios asociados a la palabra "hipoteca" o "euribor" o una marca de automóviles un anuncio para las palabras "coche" o "automóvil".

El espacio para imprimir anuncios en la primera página del listado de resultados de un buscador está limitado a 10 ítems aproximadamente como máximo. Cuando hay más de 10 anuncios asociados a una determinada palabra habrá que determinar:

- qué anuncios salen en la primer página
- qué anuncios salen en una página adjunta
- qué anuncios no saldrán

Por otra parte, aunque haya menos de 10 anuncios, también habrá que determinar en que orden quedarán situados, factor de gran importancia puesto que los usuarios suelen clicar de forma mayoritaria en los tres primeros.

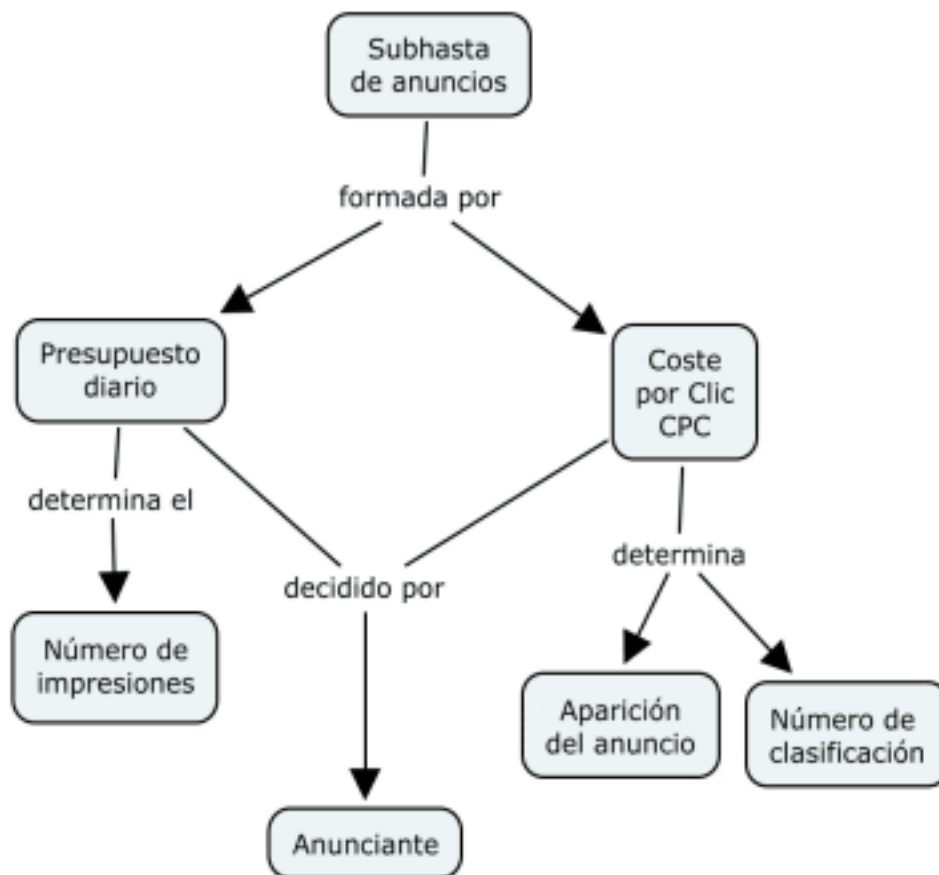
Para elegir 10 anuncios como máximo y determinar su orden, los servicios de publicidad en buscadores aplican una subasta. Los anunciantes tienen que indicar para cada anuncio que hasta que precio están dispuestos a pagar por un clic, parámetro llamado "Coste por Clic" (CPC). Cuanto más alto sea el Coste por Clic indicado, más probabilidades habrá de que el anuncio salga impreso y en buena posición y al mismo tiempo más dinero se pagará.

Resumiendo, el funcionamiento de los anuncios es el siguiente:

- Un usuario usa un buscador introduciendo algunas palabras clave
- El buscador identifica para cada palabra buscada que anuncios tiene asociados
- El buscador elige los 10 anuncios con un coste por clic y los ordena en función también del valor del coste por clic
- El buscador imprime los anuncios seleccionados
- El usuario clica sobre alguno de estos anuncios
- Finalmente el buscador facturará al anunciante el coste por todos los clics realizados

Los servicios de publicidad en buscadores buscan en todo momento la rentabilidad del anunciante. Existen multitud de opciones para facilitar a la gestión eficiente de las campañas de publicidad como por ejemplo:

- Fijar un presupuesto máximo a gastar cada día de forma que cuando se alcanza la cifra indicada deja de imprimirse el anuncio
- Sugerencia de nuevas palabras clave a partir de un listado inicial introducido por el usuario
- Sugerencia de coste por clic óptimo en función del presupuesto máximo diario o mensual



Actividad 3: Reflexione sobre las implicaciones para el anunciante del hecho de pagar por clic y no por impresión. Reflexione sobre la posibilidad de clics fraudulentos.

6.3 Palabras clave

Las palabras clave son el elemento central del marketing en buscadores. Cada anuncio tiene un listado de palabras clave asociadas que activarán su impresión y que previsiblemente serán introducidas por los usuarios de los buscadores. Una "palabra" clave puede estar formada por una o diversos términos. En función del nivel de abstracción es muy importante diferenciar entre:

- Palabras clave generales
- Palabras clave específicas (llamadas long trail)

Las palabras clave generales generan muchas más impresiones de los anuncios puesto que son usadas con mayor frecuencia por los usuarios de los buscadores. En cambio las palabras específicas generan muchas menos impresiones porque tienen una frecuencia de uso mucho más baja. No obstante, hay que tener en cuenta que el número de clics que generan las palabras específicas es proporcionalmente mucho mayor puesto que estarán más cerca de las necesidades del usuario.

- Palabras clave generales: más impresiones menos clics
- Palabras clave específicas: menos impresiones más clics

En resumen, para obtener un buen rendimiento en marketing en buscadores hay que usar aquellas palabras que sean medianamente específicas, de manera que generar un número importante de impresiones y al mismo tiempo estén lo fuertemente relacionadas con el producto o servicio promocionado para captar coincidir con las necesidades de los usuarios.

Por ejemplo, si queremos promocionar la sede web de un curso online de creación de páginas web usando software open source, no será adecuado asociar la palabra clave "curso online" puesto que será una palabra clave excesivamente general que generará muchas impresiones, pero pocos clics y además clics de baja calidad. Tampoco usaremos "open source", por las mismas razones. Tampoco usaremos "curso online paginas web open source" puesto que sería excesivamente específica. Habría que usar "curso online páginas web" y todas sus variantes.

Actividad 4: Haga una búsqueda en Google en español y otra en Yahoo en español con palabras con gran competencia. Repita la búsqueda varias veces pulsando repetidamente el botón "Buscar" y manteniendo las mismas palabras de búsqueda. Observe si los anuncios cambian. ¿Hay alguno que permanezca siempre en la primera posición? ¿podemos deducir quien está dispuesto a pagar más por clic? Repita la misma búsqueda en Google en inglés y Yahoo en inglés ¿aparecen los mismos anuncios?

7 Conclusiones

En esta unidad se ha podido ver los aspectos básicos del marketing en buscadores, del modelo del pago por clic y de la subasta de anuncios. En las próximas unidades se verán casos concretos de servicios de publicidad de buscadores donde se aplicarán estos conocimientos, servicios como Google Adwords y Yahoo! sponsored Search.

En el modelo de facturación de costes de pago por clic los anunciantes tienen la posibilidad y la responsabilidad de administrar el presupuesto de forma muy detallada. En este contexto, resulta de suma importancia establecer unos objetivos de forma clara y a continuación optimizar la campaña para su consecución. Estos aspectos serán tratados en una unidad didáctica específica para la optimización de campañas.

8 Bibliografía

Libros

Ash, Tim. Landing Page Optimization: The Definitive Guide to Testing and Tuning for Conversions. Indianapolis, IN: Wiley Publishing, 2008

Comm, Joel. The AdSense Code: What Google Never Told You About Making Money with AdSense. Garden City, NY: Morgan James, 2006

Jacobson, Howie. AdWords For Dummies. Hoboken, NJ: Wiley Publishing, 2007

Kaushik, Avinash. Web Analytics: An Hour a Day. Indianapolis, IN: Wiley Publishing, 2007

Ledford, Jerri L. Google Analytics 2.0. Indianapolis, IN: Wiley Publishing, 2007

Marshall, Perry S. Ultimate guide to Google AdWords : how to access 100 million people in 10 minutes. Irvine, Calif.: Entrepreneur Press, 2007

Mordkovich, Boris; Mordkovich, Eugene. Pay-Per-Click Search Engine Marketing Handbook: Low Cost Strategies for Attracting New Customers Using Google, MSN, Yahoo & Other Search Engines. Brooklyn, NY: MordComm, 2005

Sedes web

Click Consult. Google Adwords vs Yahoo

<http://www.clickconsult.co.uk/About-Search/Google-Adwords-vs-Yahoo.html>

Dave Manuel. Google Adwords vs. MSN Adcenter vs. Yahoo Search Marketing - The Pros and Cons of All Three.

<http://www.davemanuel.com/2007/03/27/google-adwords-vs-msn-adcenter-vs-yahoo-search-marketing-the-pros-and-cons-of-all-three/>

Functioning Form. Google vs. Yahoo! Interface Design.

<http://www.lukew.com/ff/entry.asp?189>

Google Adwords. Centro de asistencia.

<https://adwords.google.es/support/>

Google Adwords. Program Comparison

<https://adwords.google.com/select/comparison.html>

Kernelnet Informática. Esta es una comparación de los servicios de Google y Yahoo.

<http://www.kernelnet.com/content/view/236/42/>

Make. Estado de los buscadores en España: cuotas de mercado, privacidad... (Congreso OjoBuscador)

<http://www.agencia-de-publicidad.es/blog/internet/estado-de-los-buscadores-en-espana-cuotas-de-mercado-privacidad-congreso-ojobuscador/>

Ojo Buscador. Google continuará dominando en 2008, menos en Corea del Sur. Carmen López

<http://www.ojobuscador.com/2008/01/02/google-continuara-dominando-en-2008-menos-en-corea-del-sur/>

Pandia. Yahoo Search Marketing and Google AdWords compared.

<http://www.pandia.com/sw-2005/27-yahoosm-vs-adwords.html>

Profitimo. Google vs MSN vs Yahoo and SEM - By Josh Collins.

<http://www.profitimo.com/Articles/tabid/53/EntryID/820/Default.aspx>

SEOMoz. New Features in Google AdWords & Google Analytics.

<http://www.seomoz.org/blog/new-features-in-google-adwords-google-analytics>

Webfred. Las ventajas competitivas de Google Adwords sobre Yahoo Search Marketing.

<http://chocolores.wordpress.com/2007/05/21/las-ventajas-competitivas-de-google-adwords-sobre-yahoo-search-marketing/>

Your SEO Plan Pay-Per-Click (PPC) Services Comparison.

<http://www.yourseoplan.com/yahoo-google-ppc.html>

Título	Artículo V446. Vigilancia tecnológica e inteligencia competitiva para SEM-SEO
Autor	Cristòfol Rovira
Palabras clave	Vigilancia tecnológica, inteligencia competitiva, márketing en buscadores, posicionamiento web, vigilancia comercial.
Descripción	En esta unidad didáctica analizaremos con detalle en qué consiste la vigilancia tecnológica y su relación con otras disciplinas afines, como planificación estratégica, gestión del conocimiento y en especial la inteligencia competitiva a la que está unida de forma casi inseparable.
Objetivos	La orientación de esta unidad, y del módulo en general, será aplicar la vigilancia tecnológica a empresas relacionados los dos ámbitos principales del máster: SEO y SEM. La vigilancia tecnológica será especializada en este tipo de servicios y las fuentes de información, su análisis y la forma de explotar estos conocimientos será específica para las empresas dedicadas al posicionamiento y al márketing en buscadores.

Sumario

1 Introducción.....	2
2 Definiciones.....	2
3 El proceso de vigilancia tecnológica.....	4
4 Necesidades de información y factores críticos de vigilancia.....	5
5 Buscar la información necesaria para el seguimiento de los FCV.....	7
6 Análisis de la información.....	8
7 Conclusiones.....	9
8 Bibliografía.....	10

1 Introducción

En una primera aproximación la vigilancia tecnológica podemos definirla como el procedimiento sistemático de captura, análisis y explotación de información útiles para la toma de decisiones estratégicas de una empresa.

En esta unidad didáctica analizaremos con detalle en qué consiste la vigilancia tecnológica y su relación con otras disciplinas afines, como planificación estratégica, gestión del conocimiento y en especial la inteligencia competitiva a la que está unida de forma casi inseparable.

La orientación de esta unidad, y del módulo en general, será aplicar la vigilancia tecnológica a empresas relacionados los dos ámbitos principales del máster:

- Posicionamiento web
- Márqueting en buscadores

La vigilancia tecnológica será especializada en este tipo de servicios y las fuentes de información, su análisis y la forma de explotar estos conocimientos será específica para las empresas dedicadas al SEM-SEO.

Un segundo ámbito de especialización de la VT que prestaremos será el uso de recursos gratuitos:

- Se usarán principalmente las fuentes de información disponibles de forma gratuita en Internet.
- Se usarán principalmente herramientas gratuitas

Por otra parte, hay que resaltar que los buscadores también cumplen una función de herramientas para alcanzar los objetivos de cualquier tipo de vigilancia tecnológica, en especial en las fases de captura y análisis de la información.

Por tanto, los buscadores jugarán en este módulo dos roles, por una parte el tipo de empresas a las que se aplicará la VT son empresas que ofrecen servicios para los buscadores y por otra los buscadores son también una de las herramientas usadas en el proceso de la VT.

2 Definiciones

La vigilancia tecnológica "consiste en realizar de forma sistemática la captura, el análisis, la difusión y la explotación de las informaciones técnicas útiles para la supervivencia y el crecimiento de la empresa. La vigilancia debe alertar sobre cualquier innovación científica o técnica susceptible de crear oportunidades o amenazas." (Escorsa, 2001a)

La principal aplicación de la Vigilancia Tecnológica es precisamente la obtención de información de tipo técnico para la toma de decisiones del departamento de producción de una empresa. No obstante, los procesos de vigilancia se aplica también para la toma de decisiones de tipo comercial. En estos ámbitos se suelen utilizar los términos de Vigilancia Comercial, Vigilancia de la Competencia o Vigilancia del Entorno, aunque en muchas ocasiones se usa directamente el término original de Vigilancia Tecnológica que a acabado asumiendo un significado genérico.

En esta unidad y en el módulo en general se usará "Vigilancia Tecnológica" en un sentido genérico englobando tanto información de tipo técnico como de tipo comercial.

La VT está inseparablemente unida a la Inteligencia Competitiva, muchas ocasiones se usan los dos términos unidos "Vigilancia Tecnológica - Inteligencia Competitiva" (VT-IC)

Entre las dos disciplinas hay una diferencia de matiz, mientras que la VT pone el énfasis en la búsqueda y la obtención de información relevante para la toma de decisiones, la Inteligencia Competitiva se refiere al mismo proceso, pero poniendo el énfasis en la elaboración de esta información implicando a menudo la obtención de nuevas informaciones para acabar de entenderla.

"Competitive intelligence is a systematic and ethical program for gathering, analyzing, and managing any combination of Data, Information, and Knowledge concerning the Business environment in which a company operates that, when acted upon, will confer a significant Competitive advantage or enable sound decisions to be made." Society of Competitive Intelligence Professionals (SCIP)

En contextos anglohablantes se habla mucho más de "Competitive Intelligence" que de "Technology watch", en cambio en la literatura en español se usa más "Vigilancia Tecnológica" que "Inteligencia Competitiva".

En este módulo se pondrá el acento en las fases de búsqueda y obtención de la información y por tanto se usará principalmente el términos "Vigilancia Tecnológica".

Actividad 1: Compruebe hasta que punto esta afirmación es cierta buscando en Google los cuatro términos implicados (usar las comillas en el formulario de Google para hacer búsquedas con la frase completa):

- "Technology watch"
- "Competitive Intelligence"
- "Vigilancia tecnológica"
- "Inteligencia competitiva"

La VT está estrechamente relacionada con la Planificación Estratégica de una empresa que le da marco de actuación orientando su actividad al definir los factores críticos que deberá "vigilar".

En la Planificación estratégica se realizan las siguientes acciones:

1. Analizar la actividad interna y externa de la empresa
2. Identificar las Fortalezas, Debilidades, Oportunidades y Amenazas (Análisis DAFO)
3. Elaborar un plan estratégico con objetivos a corto y medio plazo
4. En función de los puntos anteriores, se definen los factores críticos de vigilancia

Por otra parte la VT se complementa con los procesos de Gestión del Conocimiento. Mientras que la VT busca y elabora la información externa a la empresa, la gestión del conocimiento está centrada en la información interna.

"La gestión del conocimiento documenta o aprovecha las experiencias de la gente de la empresa, mira hacia el pasado y busca que estos conocimientos se compartan mediante las intranets y el correo electrónico. La inteligencia [la VT] mira sobre todo al exterior de la empresa. Tiene un carácter anticipante: pretende captar señales débiles -lo que empieza a pasar, evidentemente es necesario asimilarlo lo antes posible- y, por lo tanto, busca detectar oportunidades y amenazas." (Escorsa, 2001)

Así pues, la gestión del conocimiento está pendiente de los flujos internos de información con la misión de potenciar el intercambio efectivo de información y la salvaguarda de la documentación interna de la empresa. Al mismo tiempo la vigilancia tecnológica busca y elabora información externa a la empresa que en un determinado porcentaje permitirá la toma de decisiones pasando a ser información-conocimiento interno para finalmente entrar en los procesos de la gestión del conocimiento.

3 El proceso de vigilancia tecnológica

Las fases del proceso de VT son las siguientes:

1. Identificar y analizar las necesidades de información de la empresa definiendo los factores críticos de vigilancia (FCV)
2. Buscar y obtener la información necesaria para el seguimiento de los FCV
3. Evaluar y analizar la información obtenida
4. Difundir internamente los resultados
5. Usar la información para la toma de decisiones

Estas cinco fases se ejecutan de forma continua y cíclica. A menudo las decisiones tomadas implican la aparición de nuevos FCV iniciándose un nuevo ciclo.

En los siguientes apartados aplicaremos este proceso genérico útil para cualquier tipo de empresa al caso concreto de los servicios de SEO-SEM.

4 Necesidades de información y factores críticos de vigilancia

Las principales necesidades de información externa de una empresa se pueden concretar en dos:

- Necesidad de información tecnológica: saber lo antes posible los cambios en la tecnología que podrían afectar a la empresa
- Necesidad de información comercial: conocer que hacen mis competidores y saber los cambios en mi entorno comercial en general

A partir de estos dos tipos de necesidades se pueden concretar los Factores Críticos de Vigilancia (FCV) que son factores externos a la empresa que afectan de modo crítico a su competitividad. Cada actividad e incluso cada departamento de una empresa puede tener unos FCV específicos que dependerán en gran medida de la planificación estratégica definida.

Para un buen diseño de la VT es necesario definir a los FCV de la manera más concreta posible. Por ejemplo, una empresa de SEO-SEM podría tener los siguientes FCV:

FCV tecnológicos

- Información precisa sobre qué tipo de prácticas de SEO son penalizadas por Google
- Modificaciones en el algoritmo de Page Rank
- Modificaciones en los algoritmos de ordenación de Google, Yahoo, Live...
- Novedades en las prestaciones de los servicios de márketing en buscadores
- Aparición de nuevos buscadores especializados

FCV comerciales

- ¿Cómo son los usuarios que se muestran interesados en nuestros servicios pero que todavía no son clientes?
- ¿Cómo es nuestro target?
- ¿Nuestro servicio es realmente lo que necesita nuestros clientes? ¿podemos ofrecer servicios adicionales?
- ¿Qué buscador es el líder del mercado de márketing en buscadores y en que cuota de mercado tiene?
- ¿Que comportamiento tienen los usuarios de Internet?
- ¿Qué nuevos conceptos, servicios, empresas... hay en mi sector de actividad?

FCV específicos de las campañas de SEO-SEM

- ¿En qué directorios puedo dar de alta la sede web posicionada?

- ¿Hay nuevos nombres de dominio dados de alta relacionados con la página posicionada?

Estos factores son los más generales y típicos. En cada empresa definiría algunos más en función de su planificación estratégica y sus necesidades concretas.

Actividad 2: Defina un nuevo FCV para cada uno los siguientes apartados

- FCV tecnológicos
- FCV comerciales
- FCV específicos de las campañas de SEO-SEM

Por otra parte, en el apartado de la vigilancia de la competencia y en el caso concreto de las empresas de SEM-SEO, habría que diferenciar tres tipos de competidores:

- Competencia clásica: empresas de SEO que compiten con nosotros ofreciendo los mismos o parecidos servicios
- Competencia SEO: sedes web que compiten con nosotros en el posicionamiento en buscadores con unas determinadas palabras clave
- Competencia SEM: empresas que compiten con nosotros en el posicionamiento de anuncios en buscadores con unas determinadas palabras clave

Para cada uno de estos casos habrá que definir unos FCV específicos:

Competencia clásica

- ¿Que campañas llevan nuestros principales competidores?
- ¿Qué productos y servicios ofrecen nuestros competidores?
- ¿Qué garantías ofrecen nuestros competidores?
- ¿Qué precios tienen nuestros competidores?

Competencia SEO

- ¿Qué enlaces entrantes tienen nuestros competidores?
- ¿Qué Page Rank tienen nuestros competidores?

Competencia SEM

- ¿Quiénes son nuestros competidores directos?
- ¿Qué CPM (coste por clic máximo) tienen nuestros competidores con una determinada palabra clave?

Actividad 3: Defina un nuevo FCV para cada uno los siguientes apartados

- Competencia clásica
- Competencia SEO
- Competencia SEM

5 Buscar la información necesaria para el seguimiento de los FCV

Después de identificar los FCV hay que concretar que fuentes de información vamos a usar para hacer su seguimiento y a continuación definir el procedimiento para obtener esta información de manera constante y periódica aplicando alguna de las siguientes herramientas:

- Un servicio de alerta
- Un software de monitorización de páginas web
- Un agente agregador
- Un agente de búsqueda
- Un buscador, multibuscador, metabuscador, buscador de noticias, buscador de weblogs
- Una suscripción a un canal RSS
- Un procedimiento de minería de datos
- Una base de datos bibliográfica
- Una base de datos de patentes
- Una lista de distribución
- Alguna base de datos de la web invisible

Las fuentes de información pueden ser formales o informales, en diversos tipos soporte (papel, digitales...) o sin soporte, como por ejemplo una conversación. En este módulo nos centraremos en las fuentes digitales accesibles de forma gratuita en Internet, como por ejemplo:

- Bases de datos bibliográficas de artículos de prensa especializada
- Noticias de prensa accesibles por medio de buscadores de noticias
- Patentes y normas accesibles por medio de buscadores o bases de datos de patentes
- Bases de datos y sedes web de empresas
- Estadísticas
- Blogs accesibles por medio de buscadores de blogs
- Actas de congresos, jornadas o conferencias
- Listas de distribución

Las fuentes de información informales, como conversaciones con clientes, proveedores, competidores, empleados, socios o inversores son una fuente de información muy interesante, pero hay que tener en cuenta dos elementos:

- Es necesario formalizarlas, escribirlas o grabarlas cuanto antes para evitar que se pierdan. La mínima descripción deberá contener los siguientes apartados: quien, cuando, donde y qué.
- A menudo es necesario validarlas con otra fuente

Una vez identificados los FCV y las fuentes de información pertinentes hay que asociar a cada FCV unas determinadas fuentes de información y a cada fuente

un determinado procedimiento para explotarla usando unas determinadas herramientas o servicios. Por ejemplo:

- FCV: Prácticas de SEO penalizables. Fuente de Información: Sede web de google. Herramienta: monitor de cambios
- FCV: ¿Qué productos y servicios ofrecen nuestros competidores? Fuente de Información: Sede web de la empresa competidora. Herramienta: monitor de cambios, agente agregador
- FCV: ¿Qué novedades hay en relación al PageRank de Google? Fuente de información: weblogs especializados. Herramienta: minería de datos
- FCV: ¿Qué enlaces entrantes tienen nuestros competidores? Fuente de información: buscadores. Herramienta: buscadores

Actividad 3: Elija tres FCV definidos en los anteriores ejercicios para asociarlos a unas determinadas fuentes y unas determinadas herramientas. No se preocupe si todavía le faltan conocimientos para realizar con plena garantías este ejercicio. De hecho esta asociación es uno de los principales objetivos del módulo que será tratado de nuevo en la última unidad dedicada al diseño.

6 Análisis de la información

Para tomar decisiones no es suficiente tener la información adecuada, hay que evaluarla y analizarla para seleccionar aquella que es más pertinente y presentarla de manera comprensible. No sirve de nada la obtención de grandes cantidades de datos si luego no somos capaces de procesarlos para su uso.

El resultado final más elaborado de un proceso de VT es un informe aportando datos e informaciones de diverso tipo pero de calidad, evaluados y pertinentes para la toma de una determinada decisión.

Según hemos indicado, el procedimiento concreto para elaborar este informe sería:

1. Identificación de la necesidad de información
2. Definición de los factores críticos de vigilancia (FCV) en relación a esta necesidad
3. Definición de las fuentes de información para cada FCV
4. Elección de la herramienta adecuada para sistematizar la obtención de información para cada FCV
5. Instalación y/o configuración de la herramienta de acuerdo con los requerimientos de los FCV
6. Ejecución de la herramienta y obtención de la información
7. Valoración de la información obtenida
8. Elaboración intelectual de la información valorada contrastándolos, comprendiéndolos para finalmente...
9. Redacción del informe

Este sería el procedimiento teórico. La práctica suele ser distinta por los siguientes motivos:

- No siempre es necesaria la elaboración del informe final. A menudo la propia dinámica de la empresa los hace innecesarios o inviables porque hay que tomar las decisiones de manera rápida o porque la información obtenida es lo suficientemente clara como para no dedicar recursos a la redacción de un informe.
- Las fases descritas no se ejecutan de manera completa ni de forma estricta en el orden indicado. Por ejemplo, la definición de FCV ante una necesidad de información está permanentemente abierta. En función de los datos obtenidos y su posterior elaboración se pueden detectar nuevos FCV que se incorporarán al sistema. También el uso de herramientas está abierto a ampliaciones modificando la configuración de servicios ya usados o incorporando nuevos programas informáticos.

La VT para que sea efectiva y rentable ha de tener un diseño flexible que permita la reelaboración constante de cada una de las fases sin perder el control del proceso ni de la información obtenida.

Aunque los FCV no sean muchos, es muy fácil que se genere en poco tiempo un volumen importante de información. El acceso a esta información ha de ser fácil flexible. La solución es crear un repositorio de toda la información útil de la VT aplicando las siguientes pautas:

- Toda la información codificada en un formato homogéneo (HTML, PDF...)
- Accesible por medio de algún sistema de recuperación de información, como un buscador clásico, un buscador de escritorio o una base de datos documental.

En este contexto es importante diferenciar tres momentos:

1. Revisión periódica de la información de alertas, agregadores, canales RSS, monitores, buscadores, bases de datos....
2. Almacenamiento de la información en el repositorio
3. Explotación del repositorio para la toma de decisiones y para la mejora del propio sistema

7 Conclusiones

En esta unidad didáctica se ha definido que es la vigilancia tecnológica. Se ha aplicado al caso concreto de las tareas de las campañas de posicionamiento y márketing online usando fuentes de información y herramientas disponibles de forma gratuita en Internet. La disciplina no queda agotada con lo que aquí se ha dicho puesto que su aplicación a otros ámbitos comporta el uso de otro tipo de otras fuentes y otras herramientas.

Esta unidad tiene continuidad con todas las demás del módulo, en las cuales se tratará en profundidad alguna de las fuentes de información o de las herramientas susceptibles de ser usadas en un diseño de vigilancia tecnológica.

En la última unidad del módulo se retomarán las ideas aquí expuestas para concretarlas todavía más proponiendo un diseño de un sistema de vigilancia para tareas de SEM-SEO.

8 Bibliografía

- Competitive Intelligence Dictionary <http://www.fuld.com/TIndex/IntelDict.html>
- Inteligencia Económica y Tecnológica. Guía para principiantes y profesionales http://www.madrimasd.org/informacionidi/biblioteca/publicacion/doc/15_InteligenciaEconómicaTecnológica.pdf
- Pere Escorsa Castells, Ramón Maspons. De la vigilancia tecnológica a la inteligencia competitiva. Madrid: Pearson Alhambra, 2001b
- Pere Escorsa Castells. De la vigilancia tecnológica a la inteligencia competitiva en las empresas, 2001a. http://www.uoc.es/web/esp/art/uoc/escorsa0202/escorsa0202_imp.html
- Society of Competitive Intelligence Professionals (SCIP). Language of BI <http://www.scip.org/content.cfm?itemnumber=2226&&navItemNumber=2227>
- The language of competitive business intelligence. Competitive Intelligence Glossary <http://www.quantum3.co.za/CI%20Glossary.htm>
- What Is Competitive Intelligence? <http://competitive-intelligence.mirum.net/>
- Wikipedia. Competitive intelligence http://en.wikipedia.org/wiki/Competitive_intelligence

Título	Artículo R707. Buscadores y presentación de resultados en agrupación
Autor	Mari Carmen Marcos
Palabras clave	Clustering, buscadores, agrupación, presentación de resultados, ordenación de resultados
Descripción	Se presentan las técnicas de clustering, que agrupan de forma automática los documentos obtenidos según la afinidad temática que hay entre ellos. El resultado ayuda a los usuarios a localizar temas de su interés a partir de la pregunta planteada.
Objetivos	El principal objetivo de la unidad didáctica es presentar las técnicas de clustering aplicadas a buscadores como una forma ventajosa para la presentación de los resultados de búsqueda.

Sumario

1 Introducción.....	2
2 De los listados sin orden al orden por relevancia.....	2
3 De los listados por relevancia a los clusters.....	4
4 Buscadores que presentan los resultados en clusters.....	7
5 Conclusiones.....	8
6 Bibliografía.....	9

1 Introducción

La forma en que se presentan los resultados de las búsquedas en la pantalla no es un asunto trivial. Se dedican muchos esfuerzos a la investigación en algoritmos de búsqueda y en mejorar los rankings de resultados, en cambio no hay tanta dedicación para la presentación de éstos, y por el momento apenas algunos buscadores se han atrevido a plantear formas no lineales de mostrar los resultados obtenidos.

En esta unidad se presenta la evolución de los listados de resultados para llegar a las técnicas de clustering, que presentan la ventaja de agrupar de forma automática los documentos obtenidos según la afinidad temática que hay entre ellos. El resultado ayuda a los usuarios a localizar temas de su interés a partir de la pregunta planteada.

2 De los listados sin orden al orden por relevancia

En los primeros sistemas (y actualmente también en muchos catálogos de bibliotecas) los documentos se presentaban ordenados por criterios que no eran necesariamente útiles para muchos usuarios, en el peor de los casos respondía al orden en que los registros fueron introducidos en la base de datos.



Figura 1. El catálogo de esta biblioteca ofrece los resultados de una búsqueda por materias por orden alfabético de autor, y no permite un orden cronológico

El motivo está en parte en que las bases de datos bibliográficas referenciales apenas tienen más información que la que formalmente describe el documento: autor, título, editor, ciudad, año, y poco más. Con esos datos difícilmente puede establecerse un orden de resultados más allá que el

alfabético o cronológico, pues no hay elementos para determinar la relevancia dentro del conjunto recuperado. Con la aparición de las bases de datos bibliográficas que incluyen resúmenes se podría haber hilado más fino, pero no se hizo en la mayoría de los sistemas. El cambio revolucionario vino con la llegada de las bases de datos de documentos a texto completo, donde se cuenta con un cuerpo de palabras que sí posibilita una ordenación de resultados según su relevancia.

Desde ese momento, el problema de la ordenación de los resultados se ha paliado mucho gracias a los algoritmos de recuperación de información donde cada documento tiene asignado un peso en función de cuánta es su supuesta relevancia para cada búsqueda. Así, tenemos desde los sistemas más estadísticos que comparan la frecuencia con que las palabras de búsqueda aparecen en el documento, hasta otros más sofisticados que asignan diferentes pesos según en qué parte del documento se encuentren estas palabras (el título vale más que otras partes, los primeros párrafos más que los últimos, etc.).

En este punto llegan los motores de búsqueda, que en la actualidad incorporan un criterio de ordenación que inició Google, y que consiste en dar un gran peso a la popularidad de cada sitio web recuperado, es decir, las páginas que son muy citadas por otras páginas tienen mayor probabilidad de estar en los primeros puestos que aquellas que no son tan citadas. Google tiene en cuenta además si las páginas que citan a otra son a su vez páginas muy citadas, y de esta manera se van heredando los pesos de la popularidad de las páginas web.

Actividad 1: Lea la explicación que los creadores de Google dan acerca de su forma de ordenar los resultados según su relevancia: <http://www.google.com/technology/>

Como sistema novedoso en la forma de ordenar los resultados obtenidos en buscadores podemos citar Kratia.com. Según explican sus creadores, se trata de un orden “democrático” pues está basado en los votos que los usuarios de Kratia realizan sobre los resultados obtenidos en el buscador de Microsoft, Live.com: “votando los resultados, y por lo tanto dándonos información sobre los que te gustan y los que no, podemos hacer un mejor listado, y eliminar esas web que contienen spam, que sean ilegales, o de poca calidad”.

Actividad 2: Utilice Kratia.com para realizar algunas búsquedas y revise los comentarios que hay sobre los sitios encontrados. Tenga en cuenta que se trata de un servicio aún novedoso con pocos usuarios, por lo que sólo los sitios web con muchas visitas tendrán comentarios escritos, por tanto, busque sitios muy conocidos para poder ver esta información, como en el ejemplo anterior, donde teníamos a Wikipedia.



Figura 2. Kratia busca en Live.com y ofrece a los usuarios la posibilidad de votar si el resultado le ha servido o no, así como poner un comentario sobre el sitio web

3 De los listados por relevancia a los clusters

Las llamadas técnicas de clustering tratan de agrupar objetos similares mediante algoritmos matemáticos. Estas técnicas aplicadas al ámbito de la recuperación de información consiguen crear de forma automática clasificaciones de documentos considerando ciertas similitudes en su contenido (Salton, 1979). En el caso de bases de datos documentales como los catálogos de bibliotecas, quiere decirse que se agruparán dentro de un mismo cluster aquellos documentos en los que coocuran mayor número de términos de indexación, por ejemplo.

Si a estas técnicas de agrupación se suma una presentación gráfica de su resultado, se obtiene una nueva forma de ver los documentos de las bases de datos que facilita el ojeo y por tanto la recuperación de información.

“Cluster analysis” es una expresión acuñada por Tryon (1939). Consiste en diferentes algoritmos de clasificación que organizan una cantidad de información y la convierten en conjuntos compresibles y manejables denominados clusters. Dado un conjunto de objetos (documentos) y la descripción de un conjunto A, la agrupación (clustering) debe dividir el conjunto de objetos en dos partes: los que pertenecen a A y los que no, para lo que será necesario en primer lugar establecer las características que son relevantes para describir los objetos que estarán en A (similitud intra-clustering) y en segundo lugar qué características distinguen los objetos de A de aquellos que no pertenecen a A (inter-clustering).

Otra fórmula de agrupación consiste en, a partir del conjunto de objetos y una función de similitud o distancia, dividir el conjunto de objetos en varios

subconjuntos (clusters) que revelen de la mejor manera posible la estructura del conjunto de objetos. Puede adoptar la forma de taxonomía jerárquica. En cualquier caso, se trata de que haya la máxima similitud entre los objetos que pertenecen a un cluster y la mínima entre éstos y los que no pertenecen al mismo cluster.

Para poder crear clusters los documentos se representan como vectores de términos. Hay que tener en cuenta que el tamaño de los vectores es igual al del vocabulario del conjunto recuperado. Para cada documento su vector define un punto en un espacio multidimensional. Las distancias entre puntos y su posición relativa son indicadores de la similitud entre los documentos. Para mostrar los puntos y sus relaciones al usuario es necesario reducir el número de dimensiones a una, dos o tres.

Salton (1979), en su proyecto Smart, explicaba que los documentos se representan por tantas dimensiones como número de términos de indización hay en el texto completo, generando así un espacio multidimensional que refleja todas las relaciones entre los documentos que forman la colección, pero difícil de representar mediante una visualización.

Para lograr visualizar este espacio es necesario aplicar un algoritmo de reducción y pasar a 2D o a 3D, por ejemplo el PCA (análisis de componentes principales), el MDS (escalamiento multidimensional), el análisis de cluster o el RNA (redes neuronales artificiales). En este último algoritmo, el RNA, se ha comprobado que al reducir la dimensión siempre se produce cierta distorsión. Si en lugar de reducir a 2D se reduce a 3D la distorsión es menor, pero tiene el inconveniente de que resulta más complejo de hacer (hay que pensar la metáfora muy bien) y de comprender por el usuario. Como ejemplo de reducción a dos dimensiones están los mapas autoorganizativos de Kohonen (1997) llamados SOM (Self Organizing Maps) (véase la unidad referente a visualización).

No hay mejor forma de entender un mapa de clusters que viendo y utilizando uno, como en la figura 3. Un grupo de investigación de Finlandia trabaja desde hace años en [WebSOM](http://websom.hut.fi), una herramienta que toma documentos a texto completo y mediante algoritmos de clustering presenta la colección completa en un mapa, de manera que el usuario puede saber en un solo golpe de vista qué contiene la colección, hacer un zoom en alguna de sus partes y continuar hasta localizar los documentos que más le interesan. Estos investigadores han probado con distintas colecciones de textos, entre ellas una de más de un millón de grupos de noticias Usenet en inglés.

Actividad 3: Entre en el sitio web de WEBSOM y pruebe el mapa sensible. Al clicar sobre alguna de sus zonas hará un zoom sobre ella y así sucesivamente hasta llegar a algún documento:

<http://websom.hut.fi/websom/milliondemo/html/root.html>

WEBSOM map - Million documents

Instructions

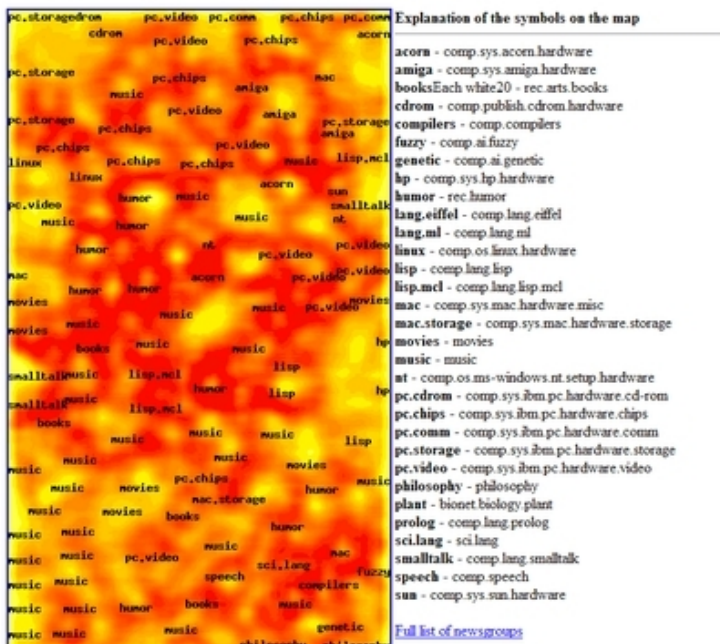


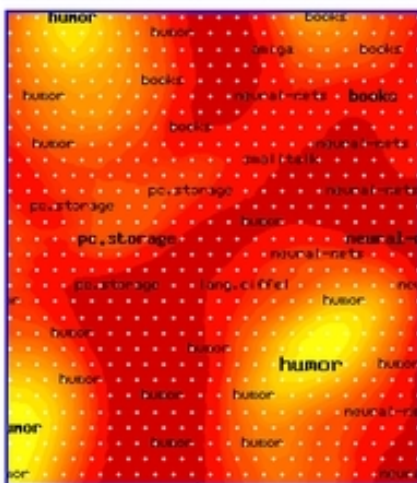
Figura 3. El mapa de clusters presenta en color amarillo los temas con más documentos, y en rojo los que menos documentos tienen.

WEBSOM zoomed map - Million documents



Click arrows

to move to neighboring areas on the map, and to move up to the overall view.



Explanation of the symbols on the map

amiga - comp.sys.amiga.hardware
books - rec.arts.books
humor - rec.humor
lang.eiffel - comp.lang.eiffel
neural-nets - neural-nets
pc.storage - comp.sys.ibm.pc.hardware.storage
smalltalk - comp.lang.smalltalk

Click any white dot to enter the node.

Figura 4. Al pulsar con el ratón se hace zoom y puede verse en detalle cada parte del mapa, sucesivamente hasta llegar a cada documento (representado con un punto blanco) e incluso leerlo, en caso de que todavía se encuentre online.

4 Buscadores que presentan los resultados en clusters

En la mayoría de los sistemas, tanto antiguos como actuales, los documentos obtenidos de una consulta se presentan sin ninguna vinculación entre sí, es más, puede haber dos documentos muy similares (mismo autor, o mismo título, o misma editorial, o misma temática) que se encontrarán alejados entre sí dentro del listado, en función del criterio de ordenación que venga dado o el que el usuario escoja (si es que puede escoger).

Trasladado a los buscadores, casi todos ellos optan por una forma secuencial de presentar los resultados, impidiendo que el usuario pueda tener una panorámica de los documentos encontrados, ni que pueda detectar si existe alguna relación entre varios de los documentos.

Existen algunos buscadores que aplican técnicas de agrupación para la presentación de los resultados, concretamente los que citaremos son multibuscadores, pues obtienen los resultados de búsquedas ejecutadas sobre los índices de otros buscadores, pero esto no importa para lo que vamos a explicar ahora.

La mejor manera de explicarlo es presentar algunos ejemplos de buscadores que han implementado técnicas de clustering para agrupar los resultados obtenidos: [iBoogie](#) y [Clusty](#) incorporan un algoritmo que tiene como objetivo crear conjuntos de documentos en función de la similitud de su contenido semántico. La forma de hacerlo, explicada de forma sencilla, consiste en agrupar los documentos en categorías según la cantidad de términos que coinciden en sus textos.

En un buscador que utilice los criterios de relevancia para ordenar los resultados, sin agruparlos por similitud, los resultados a consultas que contengan palabras polisémicas (que responden a varios significados) aparecerán mezclados entre sí en el listado. Un ejemplo: una consulta por la palabra "chile". Chile es obviamente un país, pero también es la salsa picante que usan los mexicanos para aderezar los platos.

Primero hacemos la consulta en un buscador convencional que no usa clusters, por ejemplo en Google.es, y vemos que entre los resultados obtenidos no hay ninguno que se refiera a la salsa, o si lo hay aparece mezclado con los que se refieren al país. Hacemos ahora la consulta en clusty. En la parte de la izquierda vemos cómo se forman diversas categorías (de forma totalmente automática) y se separan, entre otros, los resultados relativos a la gastronomía de aquellos que se refieren a "chile" como país (Figura 5).



Figura 5. La búsqueda "chile" en Clusty.com produce diversos clusters o grupos, uno de ellos con todos los resultados que se refieren a la salsa

El algoritmo de clustering no sólo ha creado algunos grupos con los resultados obtenidos sino que además ha puesto un título a cada uno de ellos, de manera que el usuario podrá identificar fácilmente que grupo está más cercano a sus intereses de búsqueda.

Actividad 4: Piense en algunas búsquedas con términos polisémicos, por ejemplo Jaguar (el coche y el animal), copa (para beber o como trofeo), banco (para sentarse o como entidad financiera), etc. Pruebe a buscar en Clusty.com y en iBoogie.com. Observe la ventaja de contar con clusters buscando también en los sistemas tradicionales como Google, Yahoo! o Live.

5 Conclusiones

Los buscadores han podido implementar las técnicas de recuperación de información más tradicionales de las bases de datos junto con otras gracias a contar con documentos a texto completo, donde la frecuencia estadística se puede aplicar. Con ese punto de partida, se han ido mejorando y sofisticando los algoritmos de ordenación de resultados, y en este momento todos los

buscadores cuentan con rankings de relevancia.

Uno de los puntos fuertes de Google fue llevar el conocido sistema del impacto de citas al algoritmo de ranking, con el PageRank. En cambio estos avances continúan resultando muy lineales para los usuarios, que deben revisar el listado de resultados obtenidos.

Las técnicas de clustering permiten que los documentos puedan ser agrupados en función de su similitud. Así, dos páginas web que tienen muchos términos en común podrán estar en un mismo grupo, en cambio otra que comparte menos palabras con ellas estará en un grupo diferente. El resultado de la aplicación de este algoritmo es una serie de conjuntos de documentos con afinidades entre sí. De esta forma, el usuario podrá localizar qué grupo está más cercano a sus intereses (cada grupo tiene un título que lo identifica en función de su contenido), y centrar su atención en esa parte de los documentos recuperados.

6 Bibliografía

- Kohonen, T. (1997). Exploration of very large databases by self-organizing maps. Proceedings of ICNN'97, International Conference on Neural Networks. Piscataway (NJ): IEEE Service Center, PL1-PL6.
- Larson, R. (1989). An Automatic Method of Enhancing Topical Searching for Online Catalogs based on Classification Clustering. The Annual Review of OCLC Research, July 1988-June 1989. Dublin, Ohio: OCLC Online Computer Library Center, Inc.
- Larson, R. (1991a). Classification Clustering, Probabilistic Information Retrieval and the Online Catalog. Library Quarterly, 61:2, 133-173.
- Salton, G. (1979). Progress Report on Automatic Information Retrieval. En: Second International Conference on Information Storage and Retrieval (Dallas, Texas, 27-28 September).
- Tryon, R. (1939). Cluster analysis. New York: McGraw-Hill.

Título	Artículo E665. Buscadores de mapas
Autor	Rafael Pedraza
Palabras clave	Buscadores de mapas, motores de búsqueda de mapas, Google Maps, Yahoo! Mapas, Microsoft Virtual Earth, Vía Michelin
Descripción	Se describen los buscadores de mapas más populares y se estudian sus principales características
Objetivos	El principal objetivo de la unidad didáctica es familiarizarse con los buscadores de mapas, conocer cuáles son los más importantes, cuál es su potencial, así como los principales servicios que prestan.

Sumario

1	Introducción.....	2
2	Google Maps.....	2
2.1	Mis mapas.....	7
3	Yahoo! Mapas.....	9
4	Otras opciones para la búsqueda de mapas.....	12
4.1	Microsoft Virtual Earth.....	12
4.2	Vía Michelin.....	13
5	Conclusiones.....	15

1 Introducción

En los últimos años hemos asistido a una auténtica guerra comercial en el ámbito de la recuperación de información. Los principales agentes (especialmente los buscadores) involucrados en este dominio no han parado de luchar por conseguir el predominio en la búsqueda y recuperación de cualquier tipo de información. De esta manera, el interés de estas empresas ya no se centra sólo en la recuperación de información textual, sino que progresivamente se ha ido extendiendo a otras materias, como pueden ser los blogs, imágenes, o videos. Recientemente esta batalla se ha trasladado también a la búsqueda de mapas, debido a la gran aceptación que han tenido estas aplicaciones por parte de los internautas, y la multitud de servicios con los que se relaciona (hostelería, restauración, agencias de viajes, etc.). Los buscadores de mapas se han convertido así en uno de los objetivos claves de estas empresas, que ven en ellos un negocio prometedor.

En esta lección estudiaremos los principales motores de búsqueda existentes en este momento, prestando especial atención a los dos más populares: Google Maps y Yahoo! Mapas. Para finalizar, estudiaremos otras dos iniciativas de gran interés: "Microsoft Virtual Earth", y Vía Michelin (esta última con servicios y contenidos de calidad, al menos, equiparable a los de Google y Yahoo!).

2 Google Maps

Google Maps (<http://maps.google.es/>) : es el buscador de mapas más potente de la web. Permite realizar tres **tipos de búsquedas** en mapas de todo el mundo:

Buscar en el mapa: permite encontrar una localización concreta.



Figura 1: Buscar en el mapa

Buscar negocios: nos ayuda a localizar algún tipo de negocio (hoteles, restaurantes, etc.) cercanos a una localización dada.



Figura 2: Buscar negocios

Cómo llegar: dada una dirección de origen y otra de destino, nos facilita la mejor ruta entre ambas, como si de un GPS se tratase (indicándonos en qué calles hemos de girar, qué carreteras hemos de seguir, etc.). Puede indicarse si deseamos o no hacer uso de autopistas. Además si la ruta ofrecida no es de nuestro agrado, podemos modificarla manipulando la señal/línea que representa la ruta en pantalla (en la figura 3 la línea púrpura más gruesa que puede verse), y moviéndola en el mapa hacia aquellos puntos por los que deseamos pasar (o alejándola de aquellos puntos que queremos evitar). También permite establecer nuevas rutas a partir de cada destino. Es decir, podemos fijar un recorrido con un trazado que tenga sucesivas paradas en distintos lugares. Además nos da toda la información para hacer la ruta invertida con un solo clic de ratón sobre la opción "Obtener rutas invertidas".



Figura 3: Cómo llegar

Actividad 1: Utilice Google maps para hacer una consulta de cada uno de los tipos indicados. Explore las informaciones que le ofrece este servicio.

Los mapas pueden visualizarse en tres **formatos** distintos, que podemos elegir en la parte superior derecha del mapa. La primera de las opciones se denomina sencillamente "Mapa" (figura 4) y presenta el trazado de carreteras y calles que estamos acostumbrados a encontrar en los mapas de carretera o callejeros tradicionales.



Figura 4: Vista de tipo "Mapa"

El segundo tipo de visualización se denomina "Satélite" (figura 5), y nos muestra en pantalla fotografías de satélite de la zona seleccionada.



Figura 5: Vista de tipo "Satélite"

Por último, la tercera opción se denomina "Hibrido" (figura 6), y al seleccionarla nos muestra en pantalla un mapa híbrido que combina elementos de las fotografías vía satélite (calles, plazas, etc.) con elementos de un mapa normal (introduce en las fotografías el nombre de las calles, etc.).



Figura 6: Vista de tipo "Híbrido"

El **desplazamiento a través del mapa** podemos hacerlo de dos modos:

1. Utilizando las flechas de desplazamiento que aparecen en la parte superior izquierda del mapa (flechas que indican las posiciones: arriba, abajo, izquierda y derecha).
2. Moviéndonos con el ratón, para lo cual será necesario pulsar el botón izquierdo del ratón y, manteniéndolo así pulsado, desplazar éste sobre el mapa

La **resolución** de los mapas también puede modificarse, para ello se utilizará la barra vertical que aparece bajo las flechas de desplazamiento en la esquina superior izquierda. La resolución máxima alcanzable variará de unos lugares a otros, pero suele ser tan alta que fácilmente podemos pasar de la visualización del mapa de un país a ver el callejero de cualquiera de sus ciudades o pueblos (tanto es así, que si se habilita la visión de "Satélite" es muy posible que pueda localizar su vivienda con este sistema – obsérvese que en la figura 7 es fácil discernir los coches y árboles próximos a la localización buscada, o en la figura 8, donde podemos ver con claridad la Gran Piramide de Keops en El Cairo).



Figura 7: Utilización del zoom de Google Maps sobre una localización concreta



Figura 8: Gran Pirámide de Keops (zoom de Google Maps)

Actividad 2: Elija alguna dirección o lugar de su interés y utilice las opciones de desplazamiento y zoom de Google Maps sobre ella.

Otras opciones que nos ofrece Google Maps, y que aparecen justo encima del espacio reservado para el mapa, en su esquina superior derecha, son:

- La posibilidad de convertir el mapa a una versión imprimible (clicando el enlace "Imprimir").
- La posibilidad de enviar el mapa por correo (clicando sobre el enlace "Enviar email").
- La posibilidad de insertar un mapa en nuestra página web o en un mensaje (seleccionando la opción "Enlazar con esta página").

2.1 Mis mapas

Si disponemos de una cuenta de Gmail, Google Maps nos permite acceder como usuarios registrados, personalizar un mapa o localización y guardarlo de manera que su uso sea privado o público (según nuestra preferencia). A continuación se muestra un mapa personalizado (figura 9):

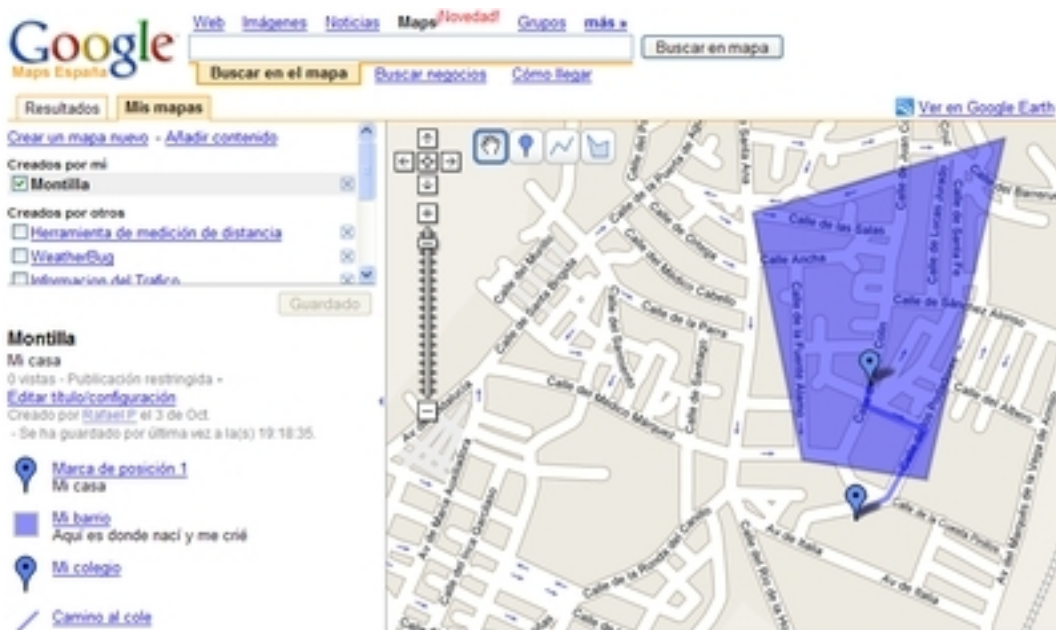


Figura 9: Ejemplo de un mapa personalizado

Este mapa representa una parte de la ciudad en la que viví durante mi niñez. En él se han incorporado los diferentes elementos que permite añadir Google Maps a un mapa, concretamente: 1. el dibujo de formas (en este caso correspondería a la zona sombreada en azul y que marcaría mi antiguo barrio); 2. Marcas de posición: son los dos punteros que aparecen en el mapa uno dentro del sombreado y otro fuera de él (y que corresponderían al lugar donde se hallaban respectivamente mi casa y colegio); 3. Líneas: que en este caso se han utilizado para unir las dos marcas de posición anteriores (simbolizan el camino que hacía de casa al colegio).

En la parte izquierda de la figura 9 vemos que aparecen todos estos elementos. Al clicar sobre cada uno de ellos se abrirá un cuadro de diálogo con su título y descripción. Además en el caso de las líneas, nos indicará la distancia en metros entre los puntos seleccionados. Estos mapas podemos guardarlos en ficheros KML (Keyhole Markup Language), un formato XML desarrollado para Google Earth que permite describir datos geográficos en tres dimensiones (título, descripción del lugar, coordenadas, etc.).

También podemos utilizar los mapas creados por otros usuarios de Google Maps (si han sido puestos a disposición pública), e incluso añadir contenidos extra a los mismos como pueden ser la posición (latitud y longitud), el tiempo,

información del tráfico, herramientas para la medición de distancias, etc. Estas herramientas pueden seleccionarse a partir del enlace "Añadir contenido" que aparece en la pestaña "Mis mapas" del margen izquierdo de la página de Google Maps (figura 11).

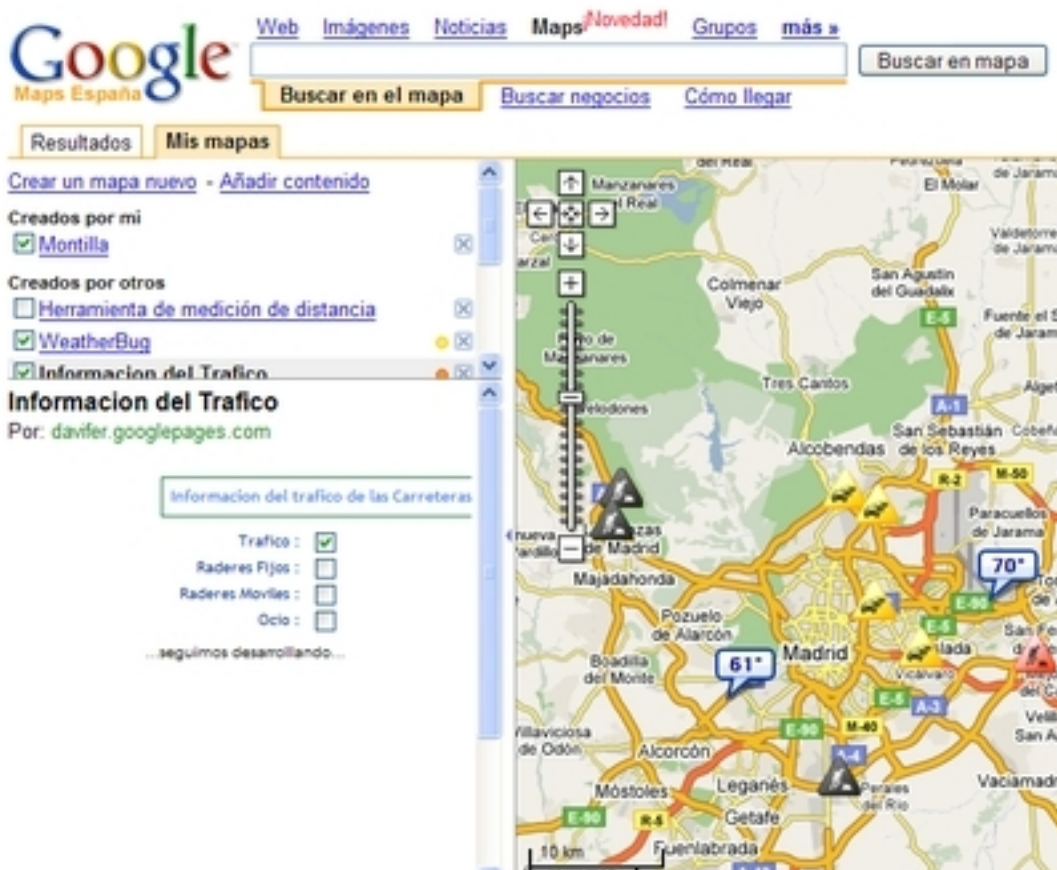


Figura 11: Con Google Maps podemos añadir contenido a los mapas

Estas son las características principales de Google Maps, no obstante ofrece muchas más, como pueden ser: la visualización de algunos lugares a través de webcams; la búsqueda (e incluso reserva) de hoteles en una ciudad; y muchas otras que pueden añadirse a este buscador para personalizarlo.

Actividad 3: Si no dispone de una cuenta de correo Gmail cree una. A continuación acceda a Google Maps como usuario registrado y cree un mapa de alguna localización de su interés, añadiéndole información personalizada. Por último, acceda a "Añadir contenido" en la pestaña "Mis mapas" y pruebe algunas de las herramientas que se le ofrecen.

3 Yahoo! Mapas

Yahoo! Mapas (<http://espanol.maps.yahoo.com/>) es un buscador dirigido principalmente a conductores. Su página inicial nos ofrece un campo de búsqueda doble (figura 12) para indicar una dirección de origen y otra de destino, aunque es posible utilizar sólo uno de ellos para buscar una localización específica. No obstante, una vez obtenida esa localización el zoom que podemos hacer sobre una zona del mapa es mucho menor que el que facilita Google Maps. Esta diferencia es especialmente notable cuando visualizamos el mapa en el modo híbrido y satélite (Yahoo! Mapas también dispone de los tres tipos de visualización que posee Google Maps: Mapa, Híbrido y Satélite).

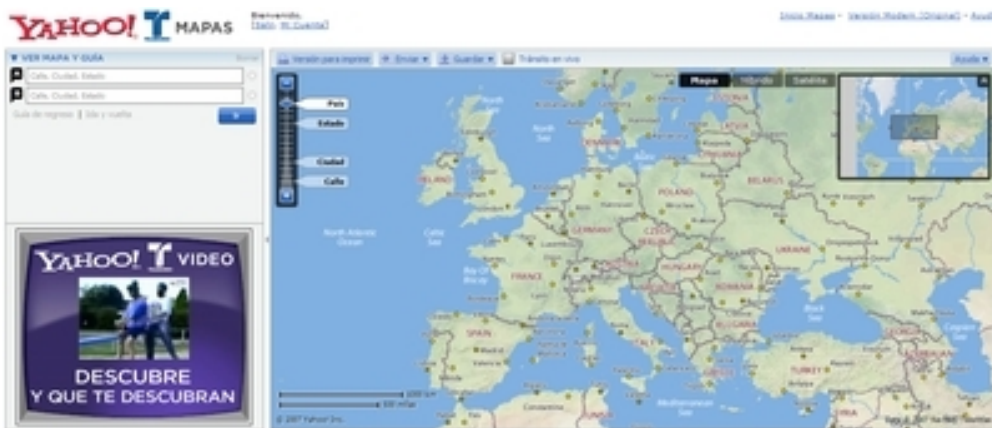


Figura 12: Página de inicio de Yahoo! Mapas

Al buscar una ruta este buscador se comporta igual que lo hace un GPS, indicándonos (figura 13) en la parte izquierda de la pantalla las calles que hemos de seguir, los puntos en los que hemos de girar, o a qué carreteras hemos de incorporarnos. Las rutas pueden establecerse entre dos o más puntos de un mapa. Además podemos solicitar una ruta de ida y otra de vuelta, o pedir que directamente nos facilite una ruta que nos valga tanto para ir como para volver. Al igual que Google Maps, nos indica la distancia entre los puntos marcados y el tiempo estimado que se tarda en recorrer la ruta.

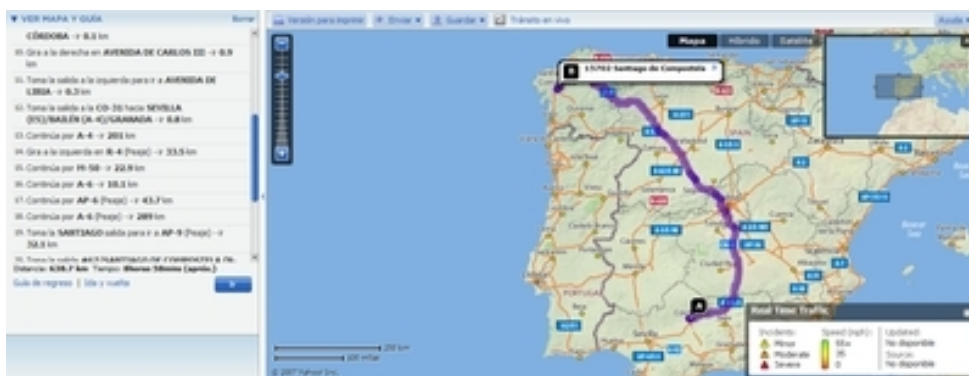


Figura 13: Especificación de una ruta con Yahoo! Mapas

Una vez escogida una ruta o localización Yahoo! Mapas nos ofrece cuatro opciones muy interesantes (figura 14): 1. La visualización del mapa de modo que sea fácilmente imprimible; 2. Enviarnos el mapa a una dirección de correo electrónico o a un coche (aunque para esto último es necesario poseer un Mercedes-Benz); 3. Guardar el mapa; 4. Visualizar el tránsito en vivo en una zona (esta opción no suele funcionar).



Figura 14: Opciones de Yahoo! Mapas

Hasta ahora se han descrito los servicios disponibles en Yahoo! Mapas sin necesidad de acceder como usuario registrado. Si disponemos de una cuenta de Yahoo! y accedemos a través de ella, el primer cambio lo apreciaremos en su página de inicio (figura 15), que ahora ofrece tres campos de búsqueda: dos para rutas (o guías de conducción), y uno específico para la obtención de mapas.



Figura 15: Inicio de Yahoo! Mapas para usuarios registrados

Un aspecto negativo de esta interfaz es la incorporación de publicidad dinámica (como podemos apreciar en la figura 15).

Al realizar una consulta en el campo "Obtener mapa", la interfaz que nos aparece también difiere de las anteriores (figura 16).



Figura 16: Resultado de una consulta en la opción "Obtener Mapa"

Como podemos ver Yahoo! sigue incorporando publicidad. En cuanto al mapa tan sólo ofrece una opción nueva: "Enlace a este mapa", que nos facilita el código fuente que hemos de introducir en nuestra página web para que dicho mapa se vea en ella.

Es sorprendente el hecho de que al utilizar Yahoo! Mapas como usuario no registrado la interfaz que se muestre sea la de mayor resolución (versión de banda ancha) y sin publicidad, mientras que al registrarnos accedemos por defecto a una interfaz de peor calidad (versión modem) y con publicidad. Por tanto, a no ser que deseemos incorporar un mapa a nuestro sitio web será más recomendable utilizar el servicio que ofrece sin necesidad de registrarnos. Si no disponemos de una conexión con banda ancha también podremos visualizar los mapas desde la página de inicio general (sin necesidad de registrarnos), seleccionando el enlace que aparece en la parte superior derecha denominado "Versión Modem (Original)".

4 Otras opciones para la búsqueda de mapas

Google Maps y Yahoo! Mapas son en este momento los servicios online más populares para la búsqueda de mapas. No obstante, también podemos encontrar otras iniciativas que merecen ser señaladas.

4.1 *Microsoft Virtual Earth*

La orientación de esta iniciativa de Microsoft (<http://www.microsoft.com/virtualearth/>) es bien distinta a las anteriores. Está dirigida a un público empresarial, y concretamente a aquellas organizaciones que quieran incorporar mapas en sus sitios web. Microsoft define a este producto como una plataforma para la localización y búsqueda de imágenes y mapas. Entre sus funcionalidades destacan la posibilidad de editar los mapas, así como la de construir mapas propios en tres dimensiones. Este software pretende dar solución a los siguientes problemas:

- Facilitar la visualización de la ubicación de ciertas localizaciones (como pueden ser tiendas, almacenes, sucursales, etc.)
- Visualización de datos sobre un mapa. Virtual Earth permite vincular un punto geográfico con un conjunto de datos de una base de datos.
- El desarrollo de mapas para portales web que faciliten determinada información a sus usuarios.
- Servicios de localización móvil: para rutas en carretera, rutas aéreas, etc.
- Virtual Earth también puede ser integrado con centralitas telefónicas, facilitando información específica sobre las llamadas (localización del lugar desde el que se realiza la llamada, etc.).
- Desarrollo de rutas comerciales para el transporte de mercancías.
- etc.

A continuación vemos (figura 17) la aplicación que ha hecho el fabricante de motos Harley-Davidson de esta tecnología. Concretamente ha instalado en su web (en la versión americana) un "Explorador de Grandes Rutas". Básicamente se trata de una aplicación que informa de manera muy completa sobre las

mejores rutas que se pueden hacer en moto en Estados Unidos. Además facilita información sobre los talleres más cercanos de Harley-Davidson a esas rutas y también sobre los hoteles/hostales en los que se puede descansar. De cada ruta ofrece una completa información: valoración de la misma (en función de la opinión de otros moteros); la mejor fecha para realizarla; frecuencia con que pueden encontrarse gasolineras; clima habitual; tipo de recorrido (con muchas curvas, recto, etc.); tráfico (abundante, moderado, etc.); desgaste de los neumáticos; lugares más interesantes para visitar; comentarios y consejos de moteros que ya han hecho la ruta; fotos; etc. En conclusión, multitud de información para todos aquellos que quieran seguir estas rutas en una motocicleta.

Actividad 4: Siga el enlace que se le muestra a continuación y ojee las funcionalidades que ofrece esta aplicación de Microsoft:

http://www.harley-davidson.com/wcm/Content/Pages/Great_Roads/great_roads.jsp?locale=en_US

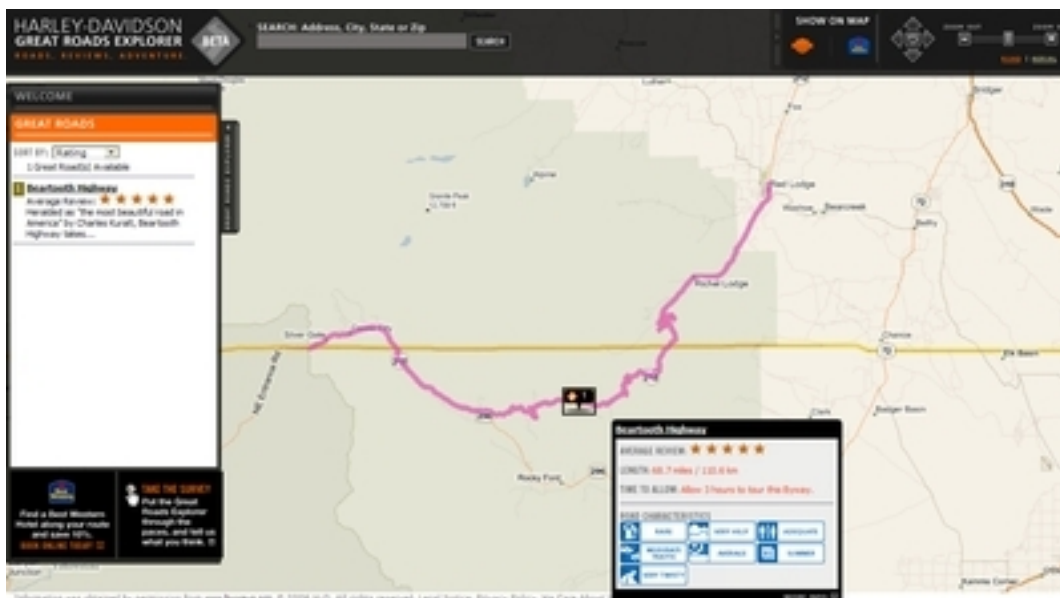


Figura 17: Microsoft Virtual Earth utilizado por Harley-Davidson

4.2 Vía Michelin

Vía Michelin (<http://www.viamichelin.es/>) es un potente buscador de mapas dirigido principalmente a conductores. Permite buscar una localización específica en mapas y planos, y una vez obtenida ofrece diferentes opciones:

- Obtener más información sobre la dirección introducida (básicamente una descripción de la ciudad e información meteorológica de la misma).
- Enviar esa dirección a nuestro GPS.
- Si somos usuarios registrados, añadir esa dirección a nuestra cuenta.

- Iniciar un itinerario o ruta con esa dirección.

Además, si lo deseamos nos facilita información muy completa de los servicios disponibles próximos a la localización introducida (figura 18): restaurantes, hoteles, parkings, radares de tráfico próximos a la zona, estaciones de cercanías y metro cercanas, e información diversa sobre el tráfico (obras, atascos, etc.). Es especialmente interesante la barra de navegación que aparece en el mapa, que además de permitir desplazarnos por el mismo y hacer zoom sobre las distintas zonas, permite activar en el mapa información diversa (estaciones de trenes, radares, etc.).



Figura 18: Servicios próximos a una localización en Vía Michelin

Este buscador también permite calcular un itinerario de viaje incluyendo nuestros parámetros personales. Podemos establecer hasta dos paradas para el cálculo de cada itinerario. Además Vía Michelin nos indicará la distancia y el tiempo estimado para realizar el recorrido. También nos hace una estimación del coste del viaje, calculada tanto en función del peaje de las autopistas, como del carburante necesario. Como resultado de la búsqueda de un itinerario obtenemos una hoja de ruta, mucho más completa que la que facilitan Google Maps y Yahoo! Mapas. En ella se detalla cuidadosamente la distancia entre cada punto de la hoja de ruta, el coste de cada peaje (si existe) y el tipo de vía por el que nos desplazaremos. Además nos permite obtener un mapa detallado de cada uno de los fragmentos de la hoja de ruta.

Toda la información así obtenida, al igual que en el caso de los mapas, podremos imprimirla, enviarla a nuestro correo, o a nuestro GPS. Además podremos crear un enlace directo desde nuestro sitio web a cualquier mapa que deseemos (siguiendo la opción "Cree un vínculo hacia este mapa"). Los mapas recuperados podrán ser visualizados a pantalla completa, mientras que las rutas se muestran en unos mapas interactivos que van mostrando poco a poco el recorrido que se habrá de realizar.

Por último, también podemos registrarnos en Vía Michelin, en cuyo caso

obtendremos permisos para guardar, comentar y compartir nuestras direcciones.

Como puede verse es un servicio muy parecido al que facilita Google y Yahoo!, aunque no dispone de visualización de localizaciones mediante fotografías vía satélite. No obstante, si nuestro objetivo es conocer los servicios, las comunicaciones, y el tráfico de un lugar al que vamos a viajar, probablemente sea ésta la herramienta más potente.

Actividad 5: Entre en Vía Michelin y realice dos consultas. La primera hágala en la sección Mapas, busque una localización concreta (preferiblemente en una gran ciudad) y a continuación active las informaciones que aparecen en el menú de navegación del mapa (en la parte superior izquierda). A continuación, localice algunos de los hoteles más próximos a su localización utilizando el menú que aparece a la izquierda del mapa. Para realizar la segunda consulta desplácese a la pestaña "Itinerarios" (en la parte superior de la pantalla). Introduzca un itinerario y analice el resultado obtenido. Observe las diferentes opciones que le facilita esta herramienta.

5 Conclusiones

En esta unidad se han estudiado las aplicaciones desarrolladas por Google, Yahoo!, Microsoft y Michelin para la búsqueda de mapas. Atendiendo a su flexibilidad podemos afirmar que las herramientas desarrolladas por Google y Yahoo! son las más potentes (siendo notoria la mejor calidad del servicio facilitado por Google Maps). No obstante, si atendemos a los contenidos mostrados en los mapas, y la versatilidad de los mismos, hemos de decir que, cuando se trata de mapas o planos (y no de imágenes o fotografías) el servicio más completo es el prestado por Vía Michelin, fruto de su gran experiencia en este sector (desde el año 2000). Microsoft no se ha mantenido al margen de la búsqueda de mapas, no obstante ha orientado su modelo de negocio de un modo completamente diferente al de sus rivales, dirigiéndose directamente a las empresas para facilitarles el software que les permita integrar un gestor de mapas con sus gestores de contenidos o plataformas. Parece que esta iniciativa ha sido acertada, y de momento grandes y reconocidas empresas como Harley-Davidson o BP están haciendo uso de su aplicación.

Junto a estas aplicaciones podemos encontrar otras muchas, normalmente dirigidas a ámbitos más restringidos (un país o grupo de países) que probablemente pronto serán eliminadas por los gigantes aquí mencionados.

Máster Online en Buscadores. Selección de unidades didácticas 2007/2008
Cristòfol Rovira , Lluís Codina , Mari-Carmen Marcos y Rafael Pedraza
1ª edición, septiembre 2008

<http://www.masterenbuscadores.com>

Editado por los autores bajo la licencia Creative Commons del tipo:
Reconocimiento-No comercial-Sin obras derivadas 2.5 España
<http://creativecommons.org/licenses/by-nc-nd/2.5/es/>

© Cristòfol Rovira , Lluís Codina , Mari-Carmen Marcos y Rafael Pedraza, 2008
ISBN: en trámite
DL: en trámite

